

STIC-Biotech/ChemLib

172599

ME

From: Li, Bao-Qun
Sent: Monday, November 28, 2005 11:55 AM
To: STIC-Biotech/ChemLib
Subject: 10,761,006

Please do the amino acid sequence homology and interference search against the nucleic acid sequence of SEQ ID NO: 1.

Bao Qun Li M.D
TC 1600
Art Unit 1648
Tel. 517-272-0904
REM, 3C18
Rm. 3D24

RECEIVED
NOV 28 2005
STIC/CHEM. DIVISION
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:14:50 ; Search time 706.5 Seconds
(without alignments)
6421.156 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907
Sequence: 1 CTCACAACTTCACCAAG.....CCTCAGCCACGAGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p_model -DEV=xlp
-Q=/cgn2.1/USFO.spool_p/US10761006/runat_01122005_113952_26869/app_query.fasta_1.3399
-DB=UniProt -QWTF=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006.qcgn 1.1199 @runat_01122005_113952_26869 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEF_TIMEOUT=120 -WARN_TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot.sprot.*

2: uniprot.trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	4130	69.9	843	2	Q69616_HPBVO	Q69616 hepatitis b
2	3745.5	63.4	832	2	Q67907_HPBVO	Q67907 hepatitis b
3	3693.5	62.5	842	2	Q69590_HPBVO	Q69590 hepatitis b
4	2824	47.8	843	2	Q7TDR5_HPBVO	Q7TDR5 hepatitis b
5	2822	47.8	843	2	Q4FD83_HPBVO	Q4FD83 hepatitis b
6	2810	47.6	840	2	Q69026_HPBVO	Q69026 hepatitis b
7	2808	47.5	843	2	Q69049_HPBVO	Q69049 hepatitis b
8	2807	47.5	843	1	DROL_HPBVR	P03157 hepatitis b
9	2806	47.5	540	2	Q97975_HPBVO	Q97975 hepatitis b
10	2806	47.5	843	2	Q5KR39_HPBVO	Q5KR39 hepatitis b
11	2804	47.5	843	2	Q7TDR3_HPBVO	Q7TDR3 hepatitis b
12	2803	47.5	843	2	Q68RQ2_HPBVO	Q68RQ2 hepatitis b
13	2803	47.5	843	2	Q4FD65_HPBVO	Q4FD65 hepatitis b
14	2802	47.4	838	2	Q81165_HPBVO	Q81165 hepatitis b
15	2802	47.4	843	2	Q42041_HPBVO	Q42041 hepatitis b
16	2802	47.4	843	2	Q5KR23_HPBVO	Q5KR23 hepatitis b

17	2801	47.4	843	2	Q80MR4_HPBVO	Q80MR4 hepatitis b
18	2800	47.4	843	2	Q5DVZ8_HPBVO	Q5DVZ8 hepatitis b
19	2799	47.4	843	2	Q5KR35_HPBVO	Q5KR35 hepatitis b
20	2799	47.4	843	2	Q4FD61_HPBVO	Q4FD61 hepatitis b
21	2798	47.4	843	2	Q91549_HPBVO	Q91549 hepatitis b
22	2798	47.4	843	2	Q80H07_HPBVO	Q80H07 hepatitis b
23	2798	47.4	843	2	Q5KR19_HPBVO	Q5KR19 hepatitis b
24	2797	47.4	843	2	Q6YLM0_HPBVO	Q6YLM0 hepatitis b
25	2796	47.3	843	2	Q80MQ8_HPBVO	Q80MQ8 hepatitis b
26	2796	47.3	843	2	Q8JVC9_HPBVO	Q8JVC9 hepatitis b
27	2795	47.3	837	2	Q5SDK1_HPBVO	Q5SDK1 hepatitis b
28	2794	47.3	843	2	Q9QNS4_HPBVO	Q9QNS4 hepatitis b
29	2794	47.3	843	2	Q9VZB5_HPBVO	Q9VZB5 hepatitis b
30	2794	47.3	843	2	Q8BC17_HPBVO	Q8BC17 hepatitis b
31	2792	47.3	843	2	Q91547_HPBVO	Q91547 hepatitis b
32	2792	47.3	843	2	Q99217_HPBVO	Q99217 hepatitis b
33	2792	47.3	843	2	Q9QM17_HPBVO	Q9QM17 hepatitis b
34	2792	47.3	843	2	Q9QM18_HPBVO	Q9QM18 hepatitis b
35	2791.5	47.3	844	2	Q4FD77_HPBVO	Q4FD77 hepatitis b
36	2791	47.2	540	2	Q97976_HPBVO	Q97976 hepatitis b
37	2791	47.2	843	2	Q39671_HPBVO	Q39671 hepatitis b
38	2791	47.2	843	2	Q81134_HPBVO	Q81134 hepatitis b
39	2791	47.2	843	2	Q8V1M4_HPBVO	Q8V1M4 hepatitis b
40	2790	47.2	843	2	Q91541_HPBVO	Q91541 hepatitis b
41	2790	47.2	843	2	Q5SDK5_HPBVO	Q5SDK5 hepatitis b
42	2790	47.2	843	2	Q5KR43_HPBVO	Q5KR43 hepatitis b
43	2790	47.2	843	2	Q5R2L9_HPBVO	Q5R2L9 hepatitis b
44	2790	47.2	843	2	Q9QMK4_HPBVO	Q9QMK4 hepatitis b
45	2789	47.2	843	2	Q9PWY8_HPBVO	Q9PWY8 hepatitis b

ALIGNMENTS

RESULT 1

Q69616_HPBVO
ID Q69616_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q69616;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name:P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype c;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnus L.O.;
RT "Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes.";
RL Virology 198:489-503(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype c;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X75656; CAAS3338.1; -; Genomic DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR000201; DNaPol_viral_N.
DR InterPro; IPR000477; RVase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.

DR Pfam; PF00078; RVT 1; 1.
 DR ProDom; PD000814; DNapol_viral_C; 1.
 DR PROSITE; P850878; RT_POL; 1.
 FT CHAIN 541 588 core antigen.
 FT NON_TER 1 1
 FT NON_TER 843 843
 SQ SEQUENCE 843 AA; 94603 MW; 934EBEC1F4E235D0 CRC64;

Alignment Scores:
 Pred. No.: 1,55e-299 Length: 843
 Score: 4130.00 Matches: 802
 Percent Similarity: 76.26% Conservative: 17
 Best Local Similarity: 74.67% Mismatches: 24
 Query Match: 69.92% Indels: 232
 DB: 2 Gaps: 1

US-10-761-006A-1 (1-3215) x Q69616_HPBVO (1-843)

QY 1 CTCACAAATTCACCAAGCTCTGTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
 Db 1 LeuHisAsnIleProThrSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
 QY 61 TGGTGGCTCCAGTCCGGACAGTAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisPheVal 40
 QY 121 AATCTTCTCGAGACTCGGGACCTCGCACCAACATGGAGAACACAAACATCAGGATTCCT 180
 Db 41 AsnLeuLeuGluAspTrpGlyProCysAsnGluHisGlyGluHisAsnIleArgIlePro 60
 QY 181 AGAACCCCTGCTGTTTACAGCGGGGTTTTCTGCTGTGACAAGAACTCTCACAAATACC 240
 Db 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
 QY 241 GCAGAGCTTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCCACGTGTC 299
 Db 81 ThrGluSerArgLeuGlyValAspPheSerGlnPheSerArgGlySerThrLysVal-Se 100
 QY 300 CTGGCCAAATTCGCAAGTCCCAACCTCCAATCAGTCCCAACCTCTGCTCTCCAATTT 359
 Db 100 TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 120
 QY 360 GTCTGGCTATCGCTGGATGTGTCTGCGCGGTTTTATCATATTCCTTCTCATCTGCTGC 419
 Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 140
 QY 420 TATGCTCATCTCTTGTGTGTCTCTGACTTACCAAGGTATGTGCGGTTTGCTCTC 479
 Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGAACATCAACACACAGCACGGGGCCATGCAAGACCTGCAGACTCTCTGTCTC 539
 Db 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 180
 QY 540 AAGAAACTCTAGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
 Db 180 rArgAsnLeuTyrValSerLeuMetLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
 QY 600 GTATTCCATCCATCATCTGGGCTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
 Db 200 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTTCCTCTGGCTCAGTTACTAGTCCATTTGTTTCAGTGGTTCTGAGGGCTTCCCCCA 719
 Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGTTATTTGGGGCGAAGTCTGTACAACATCT 779
 Db 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
 Db 260 uGluSerLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280

QY 840 TAATAAAACCAACAGCTTGGGGCTACTCCCTTAACCTTCACTGGGATATGAATTGGAAGTTG 899
 Db 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
 QY 900 GGGTACTTTACCGCAGGAACATATTTGTACTATAAACTCAAGCAATGTTTTTCGAAAACTGCC 959
 Db 300 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAATTTGGGCTCTTTTGGGCTTTGTC 1019
 Db 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTCTTTTACAAATGTGGCTATCTCTGCCCTTGTATGCTTTTATATGATGATATCAATC 1079
 Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 360
 QY 1080 TAAGCAGGCTTTCACTTCTCGCCCAACTTACAAGCCCTTCTGTGTAAACAATATCTGAA 1139
 Db 360 aLysGlnAlaPheThrPheSerSerThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 380
 QY 1140 CCTTTTACCCCTGTCGCGCAACGCTCGCTCTGCCAAGTGTGTGCTGACGCCAACCC 1199
 Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTCGATGGGGCTTTGGCCATAGGCATCAGCGCATGGCTGGAACCTTTCTGGCTCCTCT 1259
 Db 400 oThrGlyTrpGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCCGATCCATCTACTGCGGAACCTCTAGCAGCTTTGTTTGTCTGCACGCCGCTCTGGAGCAA 1319
 Db 420 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 440
 QY 1320 ACTTATCGAACCGCAACTCTGTGTGTCTCTCTCGGAATAACACCTCTTCTTCCATGGCT 1379
 Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
 QY 1380 GCTAGGCTGTCTGCAACTGATCTCTGCGGGAGCTCTTGTGTCTACGCTCCCTCGGC 1439
 Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
 QY 1440 GCTGAATCCCGCGAGCACCCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
 Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuAr 500
 QY 1500 TCTGCGGTTTCCGGCGACACCGGCGCACCTCTCTTTACGCGGCTCCCGTATGTGCC 1559
 Db 500 gLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
 QY 1560 TTCTCATCTGCGGACCGTGTGCACTTGCCTTCACTCTGCACTGCGATGGAGACCACC 1619
 Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
 QY 1620 GTGAACGACGACCGCAGGTCTTGCCCAAGGTCTTATATAAGAGACTCTTGGACTCTCAGCA 1679
 Db 540 o----- 540
 QY 1680 ATGTCAACGACCGACTTGGAGCATACTTCAAGACTGTGTGTTTAAAGACTGGGAGGAG 1739
 Db 540 ----- 540
 QY 1740 TTGGGGGAGAGATTAGGTAAAGATTATGTACTAGGAGGCTGTAGGCATAAATTTGGTC 1799
 Db 540 ----- 540
 QY 1800 TGTTCACGACGACCATGCAACTTTTCTCTCTGCTTAATCATCTCATGTTCATGTCCTA 1859
 Db 540 ----- 540
 QY 1860 CTGTTCAAGCCCTCAAGCTGTGCTTGGGTGGCTTTGGGACATGGACATTGACCCGTATA 1919
 Db 540 ----- 540


```
QY 1920 AAGAAATTGGAGCATCTGCTGAGTACTCTCTTTTGGCCCTTCTGACTCTTTCCGCTCTA 1979
Db 540 -----
QY 1980 TTCGAGATCTCTCGACACCGCCTCTGCTCTGTATCGGAGGCTTAGAGTCTCCGGAAC 2039
Db 540 -----
QY 2040 ATTGTCGCCTCACCATACAGCACTCAGCAAGCTATTTTGTGTGGGTGAGTGTGATGA 2099
Db 540 -----
QY 2100 ATCTGGCCACTGGGTGGAGTAATTTGGAGATCCAGCATCCAGGAATTAGTAGTCA 2159
Db 540 -----
QY 2160 GCTATGTCAAGTTAATATGGGCTCAAACTCAGACAAATATTGTGGTTTCACATTTCC 2219
Db 540 -----
QY 2220 GTCTTACTTTTGGAGAGAACTGTCTTCTGAGTACTTGGTATCTTTTGGAGTGTGGATTC 2279
Db 540 -----
QY 2280 GCATCTCCTACCGTTACAGACCACCAATGCCCTATCTTATCAACTTCCGGAACTA 2339
Db 541 -----Met---ProLeuSerTyrGlnHisPheArgLysLeu 551
QY 2340 CTGTTGTTAGACGAGGAGGCTCCCTAGAGAACTCCCTCGCTCGCCGAGACGAA 2399
Db 552 LeuLeuLeuAspValGluAlaGlyProLeuGluGluLeuProArgLeuAlaAspGlu 571
QY 2400 GGTCTCAATCCCGCGTCGAGAGATCTCAATCTCGGGAATCTCAACGTTAGTATTCCT 2459
Db 572 GlyLeuAsnHisArgValAlaGluAspLeuAsnLeuGlyAsnProAsnValSerIlePro 591
QY 2460 TGGACTCATAAAGTGGGAACTTTACTGGGCTTTTATCTTCTTACTGACTGCTTTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuTyrSerSerThrValProValPheAsn 611
QY 2520 CCCAGTGGCAAAATCTCTCTCTCTCTCTCACATTTTACAGAGGACATTTAATAAGA 2579
Db 612 ProGluTrpGlnThrProSerPheProAspIleHisLeuGluAspIleIleAsnArg 631
QY 2580 TGTCAACAATATGTGGGCTCTTACAGTTAATGAAAGAAAGATTAATAATTAATG 2639
Db 632 CysGlnGlnTyrValGlyProLeuThrValAsnGluLysArgLeuLysLeuIleMet 651
QY 2640 CTGCTAGTGTATTCCTTAACCTTACTAAATATTTGGCTTTAGACAAAGGCATTAACCG 2699
Db 652 ProAlaArgPheTyrProAsnLeuThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
QY 2700 TATTATCTGACATGCGATTAACTTACTTCAAACTAGGCATTTTACATCTCTG 2759
Db 672 TyrTyrProGluHisAlaValAsnHisTyrPheLysThrArgHisTyrLeuHisThrLeu 691
QY 2760 TGGAGGCTGCGATTCTATATAAGAGAGAACTACACGACGCTCTATTTGTGGGTCA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrArgSerAlaSerPheCysGlySer 711
QY 2820 CCATATCTTGGGAACAGAGCTACAGATCGGAGGTGGTCTTCCAAACCTCGACAAAG 2879
Db 712 ProTyrSerTrpGluGlnGluLeuGlnHisGlyArgLeuValPheGlnThrSerGluArg 731
QY 2880 CATGGGAGCAATCTTGTCTGTTCCTCAATCTCTGGGATCTTTCCTGATCACAGTTGGA 2939
Db 732 HisGlyAspGlnSerPheCysSerGlnSerGlyIleLeuSerArgSerProValGly 751
QY 2940 CCCTCGTTCGAGGCAACTCAACAATCCAGATTTGGACTTCAACCCCAACAGGATCA 2999
Db 752 ProCysIleArgSerGlnLeuLysGlnSerArgLeuGlnProGlnGlnGlySer 771
QY 3000 CTGGCCAGGACCAATCAGGTAGGAGTGGGAGCATTCGGGGCCAGGGTTTCACCCCACCACA 3059
```

```
Db 772 MetAlaArgGlySerGlyArgSerGlySerGlySerIleArgAlaArgValHisProThrThr 791
QY 3060 CGCGCGTCTTTTGGGGGAGCCCTCAGGCTCAGGCATATTATTCACACAGTGCACAGC 3119
Db 792 ArgArgSerPheGlyValGluProSerGlyThrArgHisIleAspAsnThrAlaSerSer 811
QY 3120 ACTCTCTCTCTCTCCACCAATCGCAGTCAGAGACAGACCTTCTCCATCTCTCCACC 3179
Db 812 SerSerSerCysLeuHisGlnSerAlaValArgLysThrAlaTyrAlaHisLeuSerThr 831
QY 3180 TCTAAGAGACAGTCTATCTCTCAGCCACGACGAGTGGAA 3215
Db 832 SerLysArgGlnSerSerSerGlyHisAlaValGlu 843

RESULT 2
Q67907 HPBV0
ID Q67907 HPBV0 PRELIMINARY; PRT; 832 AA.
AC Q67907_
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE DNA polymerase (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sub-type ayw;
RX MEDLINE=94079539; PubMed=8257295;
RA Preisler-Adams S., Schlayer M.J., Peters T., Hettler F., Gerok W.,
RA Rasenack J.
RT "Sequence analysis of hepatitis B virus DNA in immunologically
RT negative infection.";
RL Arch. Virol. 133:385-396 (1993).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X72703; CAAS1254.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR01462; DNAPol_viral_C.
DR InterPro; IPR00201; DNAPol_viral_N.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS50878; RT_POL; 1.
FT NON_TER 1
FT NON_TER 832
SQ SEQUENCE 832 AA; 93864 MW; EC6872097F04BAAP CRC64;

Alignment Scores:
Pred. No.: 8,42e-271 Length: 832
Score: 3745.50 Matches: 726
Percent Similarity: 72.16% Conservative: 49
Best Local Similarity: 67.60% Mismatches: 57
Query Match: 63.41% Indels: 243
DB: 2 Gaps: 2

US-10-761-006a-1 (1-3215) x Q67907_HPVB0 (1-832)
QY 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGCTGAGGCTATATTTCTCTGC 60
Db 1 LeuHisAsnPheProProAsnSerAlaArgSerGlnGlyGluArgProValPheProCys 20
QY 61 TGTGGCTCCAGTTCGGACAGCTAACCTGTTCCGACTTCCGCTCTCCCATATCGTC 120
Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 40
```

QY	121	AATCTCTCGAGACTGGGGACCTCGACCGAAACATGGAGAACACAACTCAGGATCTCT	180
Db	41	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro	60
QY	181	AGGACCCCTGCTCGTGTACAGCGGSGTTTTCTCGTTGACAGAAATCCTCACATACC	240
Db	61	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	80
QY	241	GCAGACTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGTT	299
Db	81	AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se	100
QY	300	CTGGCCAAATTCGGAGTCCCAACCTCCAACTCACATCACCAACCTTGTCTCTCAATTT	359
Db	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
QY	360	GTCTCTGGCTATCGCTGATGTGTCGGCGGTTTTATCATATTCCTCTTCATCCTGCTGC	419
Db	120	userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe	140
QY	420	TATGCTCATCTCTTGTGTGTTCTTCTGCACTTACCAAGGTATGTTGCCGTTTGCTTC	479
Db	140	rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe	160
QY	480	TACTTCCAGAAACATCAACACACAGCACGGGGCCATGCAAGACCTGCACGACTCCTGCTC	539
Db	160	rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe	180
QY	540	RAGGAACCTAGCTTTCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACTGCATTT	599
Db	180	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe	200
QY	600	GTATTCCTCATCCATCATCTGGGCTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC	659
Db	200	uTyrSerHisProIleLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerPr	220
QY	660	GTTCCTCTCGGCTCAGTTACTAGTCCCAATTTGTTCCAGTGGTTCGTAGGGCTTTCGCCA	719
Db	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
QY	720	CTGTTTGGCTTCAGTTATATGATGATGTTGTTGGGGCGAAGTCTGTACACATCT	779
Db	240	sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe	260
QY	780	TGAGTCCCTTTTACCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC	839
Db	260	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
QY	840	TAATAAAACCAACGTTGGGCTTACTCCCTTAACCTCATGGGATATGTAATTGGAAGTTG	899
Db	280	oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy	300
QY	900	GGGTACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTCGAAAACCTGCC	959
Db	300	rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr	320
QY	960	TGTAATAGACCTATTGATTGGAAGTATGTCTCAAGAAATTTGGGGTCTTTTGGGCTTTCG	1019
Db	320	oIleAsnAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
QY	1020	TGCCCCCTTTTACAAATGTGGCTATCTCGCTTGTATGCCCTTTTATATGATGTATACAATC	1079
Db	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
QY	1080	TAAGCAGGCTTTCACATTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA	1139
Db	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	380
QY	1140	CTTTTACCCCGTTCGCCGCAACGGTCTCTCGCAAGTGTTCCTGACGCAACCC	1199
Db	380	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	400
QY	1200	CACTGGATGGGGCTTTGGCCATAGGCCATAGGCCATGGCTGGAACCTTTCTGGCTCCTCT	1259
Db	400	oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProIle	420
QY	1260	GCGGATCCATCTACTCGGAACCTCTAGCAGCTTGTGTTGCTCGCAGCCGCTCTGGAGCAAA	1319
Db	420	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs	440
QY	1320	ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACCTCTCTTTCCATGGCT	1379
Db	440	nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpIle	460
QY	1380	GCTAGGTGTGCTGCCAACTGGATCCTCGCGGGAGCGTCTTGTGCTAGTCCCGTCGGC	1439
Db	460	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	480
QY	1440	GCTGAATCCCGCGGACGACCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTTCTCA	1499
Db	480	aleuAsnProAlaAspAspProSerArgLysArgLeuGlyIlePheArgProLeuLeuAr	500
QY	1500	TCTGCCGTTTCGGCGCGACACGCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC	1559
Db	500	gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr	520
QY	1560	TTCCTCATCTGCGGACCGGTGTGCATTCGCTTTCACCTCTGCACTGCGCATGGAGACCA	1619
Db	520	oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	540
QY	1620	GTGAACGCAACCCAGGTCTTGCCCAAGGTCTTATATATAGAGACTCTTGGACTCTCAGCA	1679
Db	540	o-----	540
QY	1680	ARGTCAACGACCGACTTTGAGGCATATCTTCAAAGACTGTGTGTTTAAAGACTGGGAGG	1739
Db	540	-----	540
QY	1740	TTGGGGAGGAGATTAGGTTAAAGATTATGTACTAGGAGGCTGTAGGCATAAATTTGGTC	1799
Db	540	-----	540
QY	1800	TGTTCAACGACCACTGCAATTTTCTCTCTGCTGCTAATCATCTCATGTTCATGTCCTA	1859
Db	540	-----	540
QY	1860	CTGTTCAAGCCTCCAAGCTGTGCCCTTGGGTGGCTTTTGGGACATGCACATTTGACCCGTATA	1919
Db	540	-----	540
QY	1920	AAGAAATTTGGAGCATCTGCTGAGTTACTCTCTTTTTCCTTCTGACTTCTTTCCGTCTA	1979
Db	540	-----	540
QY	1980	TTGAGATCTCTCTGACACCGCTCTGCTGCTATCGGAGGCGCTTAGAGTCTCCGGAAC	2039
Db	540	-----	540
QY	2040	ATTGTTGCGCTCACCATACAGCACTCAGCAAGCTATTTTGTGTTGGGTGAGTTGATGA	2099
Db	540	-----	540
QY	2100	ATCTGGCCACTGGGTGGGAAGTAATTTGGAAGATCCAGCATCCAGGGAATTAGTAGTCA	2159
Db	540	-----	540
QY	2160	GCTATGTCAACGTTAATATGGSCCTAAAACTCAGACAAATATTGTGGTTTCACATTTCT	2219
Db	540	-----	540
QY	2220	GTCTTACTTTTGGGAAGAGAACTGTTCTTGTAGTACTTGTGTATCTTTTGGAGTGTGATT	2279
Db	540	-----	540
QY	2280	GCACCTCTACCGCTTTCAGACCAACCAATGCCCTATCTTATCAACACTTCCGGAACATA	2339

Db 541 -----Met-----ProLeuSerTyrGlnHisPheArgLysLeu 551
QY 2340 CTGTTGTTAGACGACGAGGAGCTCCCTAGAGAAGAACTCCCTCGCTCGCAGACGAA 2399
Db 552 LeuLeuLeuAspGluAlaGlyProLeuGluGluGluLeuProArgLeuAlaAspGlu 571
QY 2400 GGTCTCAATCGCGCTCGCAGAGATCTCAATCTCGGGAATCTCAAGTTAGTATCTCT 2459
Db 572 AspLeuAsnArgValAlaGluAspLeuAsnLeuGlyAsnLeuAsnValSerIlePro 591
QY 2460 TGGACTCATAGGTGGGAACCTTACTGGGCTTTATTCTCTACTGTACCTGCTCTTAAT 2519
Db 592 TrpThrHisLysValGlyAsnPheThrGlyLeuSerSerThrValProValPheAsn 611
QY 2520 CCGAGTGGCAAAATCTCTCTCTCTCAATCTTACAAGAGGACATTAATATAGA 2579
Db 612 ProHisTrpLysThrProSerPheProAsnIleHisLeuHisGlnAspIleLysLys 631
QY 2580 TGTCAACATATGTGGGCTCTTACAGTTAATGAAAAGAGATTAATAATATATG 2639
Db 632 CysGluGlnPheValGlyProLeuThrValAsnGluLysArgArgLeuGlnLeuIleMet 651
QY 2640 CCTGCTAGGTTTATCTTACCTTACTAATAATATTTGCTTACAGCAAGGCATTAAACCG 2699
Db 652 ProAlaArgPheTyrProAsnValThrLysTyrLeuProLeuAspLysGlyIleLysPro 671
QY 2700 TATTATCTGAACATGCAAGTTAATCAATCTTCTCAAACTAGGCAATTTTACATCTCTG 2759
Db 672 TyrTyrProGluHisLeuValAsnHisTyrPheGlnThrArgHisTyrLeuHisThrLeu 691
QY 2760 TGGAGCTGGCATCTTATATAGAGAGAACTACACGCGCGCTCATTTTGTGGGTCA 2819
Db 692 TrpLysAlaGlyIleLeuTyrLysArgGluThrThrArgSerAlaSerPheCysGlySer 711
QY 2820 CCATATCTTGGGAACAGAGCTACACATGGGAGGTGGTCTTCCAACTCGACAAAG 2879
Db 712 ProTyrSerTrpGluGlnGluLeuGln----- 720
QY 2880 CATGGGAGCAATCTGCTGTTCCCAATCTCTGGGATCTTCCCGATCACCAGTTGGA 2939
Db 721 HisGlyAlaGluSerPheHisGlnGlnSerSerGlyIleLeuSerArgProProValGly 740
QY 2940 CCCTGCTTGGAGCAACTCAACAATCAGATGGGACTTCAACCCCAACAGGATCA 2999
Db 741 SerSerLeuGlnSerLysHisArgLysSerArgLeuGlyLeuGlnSerGlnGlnHis 760
QY 3000 CTGGCCAGGCAATCAGTAGAGTGGGAGCATTCGGCCAGGTTTACCCACCACCA 3059
Db 761 LeuAlaArgArgGlnGlnGlyArgSerTrpSerIleArgAlaGlyIleHisProThrAla 780
QY 3060 CGGCGGTCTTTTGGGGGGAGCCCTCAGGCTCAGGCAATTTGACAAACAGTCCAGCAGC 3119
Db 781 ArgArgProPheGlyValGluProSerGlySerGlyHisAsnThrAsnLeuAlaSerLys 800
QY 3120 ACCTCTCTGCTCCCAATCGCAGTCCAGGAGACAGCTTACTCCCATCTCTCCACC 3179
Db 801 SerAlaSerCysIleTyrGlnSerProValArgLysAlaAlaTyrProAlaValSerThr 820
QY 3180 TCTAAGAGACAGTCACTCTCAGGCCAGCAGTGGAA 3215
Db 821 PheGluLysHisSerSerSerGlyHisAlaValGlu 832
RESULT 3
Q69590_HPBVO
ID Q69590_HPBVO PRELIMINARY; PRT; 842 AA.
AC Q69590;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polymerase (Fragment).
GN Name=P;
OS Hepatitis B virus.

OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RX MEDLINE=94120723; PubMed=8291231;
RA Norder H., Courouce A.M., Magnius L.O.;
RT "complete genomes, phylogenetic relatedness, and structural proteins
of six strains of the hepatitis B virus, four of which represent two
new genotypes.";
RL Virology 198:489-503 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype e;
RA Norder H.M.L.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X75657; CAA53339.1; -: Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_N.
DR InterPro; IPR000201; DNAPol_viral_N.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_POL; 1.
FT CHAIN <1> >842 Potential.
FT NON_TER 1
FT NON_TER 842
FT SEQUENCE 842 AA; 94812 MW; 45C3CD71CF5A4FE7 CRC64;
SQ
Alignment Scores:
Pred. No.: 6,51e-267 Length: 842
Score: 3693.50 Matches: 719
Percent Similarity: 71.85% Conservative: 52
Best Local Similarity: 67.01% Mismatches: 70
Query Match: 62.53% Indels: 233
DB: 2 Gaps: 2
US-10-761-006A-1 (1-3215) x Q69590_HPBVO (1-842)
QY 4 CACAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCTATATTTCTCTGCTGG 63
Db 2 HisAsnIleProProSerSerAlaGlySerLysArgProValPheSerCysTrp 21
QY 64 TGCTCCAGTTCCGGAAACAGTAAACCTGTTCGACTACTCTCTCCATATCGTCAAT 123
Db 22 TrpLeuGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisLeuValAsn 41
QY 124 CTCTCTCAGGACTGGGGACCTCGACGAACTGGAGAACACCAATCAGATTCCTAGG 183
Db 42 LeuLeuGluAspTrpGlyProCysThrGluHisGlyLysHisIleArgIleProArg 61
QY 184 ACCCTCTGCTGTTTACAGCGGGGTTTTCTCGTTCAGCAAGATCTCTCAATACCGCA 243
Db 62 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrAla 81
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGCTCTCTG 302
Db 82 GluSerArgLeuValValAspPheSerGlnPheSerArgLysSerArgVal-SerTr 101
QY 303 GCCAAATTCGAGTCCCAACCTCCCAATCACCACCACTCTGTCTCTCAATTTGTC 362
Db 101 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 121
QY 363 CTGGCTATCGCTGATGTGTCTGGGGGTTTTTATTCATATTTCTCTTCTGCTGCTAT 422

Db		121	rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAlaMe	141	Db	uAsnProAlaAspProSerArgGlyArgLeuGlyIleTyrArgProLeuLeuArgIle	501	
QY	GCCTCATCTCTTGTGGTCTCTCGACATACCAGGTATGTTGCCGGTGTCTCTCTAC	482		142	QY	GCCTGTTCCGGCGCAGCACCGGGCGCACCCTCTCTTTACCGGGTCTCCCGTATGTGCCTTC	1562	
Db		141	tProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaAlaArgLeuSerSerAs	161	Db	uProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe	521	
QY	TTCCAGGAACATCAACCCAGCAGCGGGCCCATGCAAGACCTGCGACGATCTCTGCTCAAG	542		162	QY	TCATCTGCGGAGCCGCTGTGCACTTCGCTTACCTCTGACGCTGCGATGAGAGACCGGTG	1622	
Db		161	nSerArgIleIleAsnHisGlnTyrGlyThrLeuProAsnLeuHisAspSerCysSerAr	181	Db	rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTyrArgProPro	540	
QY	GAACCTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGACCGGAACCTGCACTTGT	602		182	QY	RACGCACGCCAGGTCTTGCCCAAGGCTTATATATAGAGGACTCTTGGACTCTCAGCAATG	1682	
Db		181	gAsnLeuTyrValSerLeuMetLeuLeuPheLysThrPheGlyArgLysLeuHisLeuTy	201	Db		540	540
QY	TTCCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTT	662		202	QY	TCACGACCGACCTTGAGGCATACCTCAAGACTGTGTGTTTAAAGACTGGGAGGAGTTG	1742	
Db		201	rSerHisProIleIleMetGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	221	Db		540	540
QY	TTCTCTGGCTCAGTTTACTAGTCCATTTGTTCACTGGTTCGTAGGGCTTTCCCCCACTG	722		222	QY	GGGAGGAGATTAGGTTAAAGATTTATGTACTAGGAGGCTGTAGGCATATAATTTGCTCT	1802	
Db		221	eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy	241	Db		540	540
QY	TTTGGCTTTCACTATATGATGATGCTGTATTTGGGGCGAAGTCTGTACACATCTTGA	782		242	QY	TCACCAGCACCATGCAACTTTTCTCTCTGCTTAATCATCTCATGTTTCATGTCCTACTG	1862	
Db		241	sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	261	Db		540	540
QY	GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCTAA	842		262	QY	TTCAAGCCTCCAAGCTGTGCTTGGGTGGCTTTTGGGACATGACATTTGACCCGTATAAG	1922	
Db		261	uSerLeuTyrThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs	281	Db		540	540
QY	TAAACCAACGTTGGGCTACTCCCTTAACCTTATGGGATATGTAATTTGGAAGTTGGGG	902		282	QY	AATTTGGAGCATCTGCTGAGTTACTCTCTTTTTTTCGCTTCTGACTTCTTCCGCTCTATT	1982	
Db		281	nLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyrGl	301	Db		540	540
QY	TACTTTACCGCAGGACATATGTACTAAACTCAAGCAATGTTTCGAAAACCTGCTGT	962		302	QY	GAGATCTCTCGACACCGCCTCTGCTCTGTATCGGGAGGCCCTTAGAGTCTCCGGAACAT	2042	
Db		301	ySerLeuProGlnGluHisIleGlnLysIleLysAspCysPheArgLysLeuProVa	321	Db		540	540
QY	AAATAGACTATTGATTGAAAGTATGTCAAAGATTGTTGGGTCTTTTGGGCTTTGCTGC	1022		322	QY	GTTCGCTCTACCATACAGCACCTCAGCAAGCTATTTTTGTGTTGGGTGAGTTGATGAATC	2102	
Db		321	lAsnArgProIleAspTyrLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	341	Db		540	540
QY	CCCTTTTACAAATGTGGCTTCTCTGCTTGTATGCTTTATATGATGATATCAATCTAA	1082		342	QY	TGGCCACCTGGGTGGGAGTAATTTTGGAGATCCAGCATCCAGGGAATTTAGTAGTCAGCT	2162	
Db		341	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysThrGlnSerLy	361	Db		540	540
QY	GCAGGCTTTCACCTTCTGCCCACTTACAAGCCTTCTGTGTAAACAATATCTGAACCT	1142		362	QY	ATGTCAACGTTAATATGCGCCTTAAACTCAGACAAATATTGTTGTTTCACATTTCCCTGTC	2222	
Db		361	sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe	381	Db		540	540
QY	TTACCCCGTTTGGCCGCAACGGTCCGGTCTCTGCCAAGTGTTCCTGACGCAACCCAC	1202		382	QY	TTACTTTTGGAGAGAACTGTTTCTTGAGTACTTGTGTATCTTTTGGAGTGTGGATTTCGA	2282	
Db		381	uTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrProTh	401	Db		540	540
QY	TGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTCTGCTGCTCTGCTGCC	1262		402	QY	CTTCTACCGCTTACAGACACCAAAATGCCCCTATCTTATCAACACTTCCGGAACACTACTG	2342	
Db		401	rGlyTyrGlyLeuAlaIleGlyIleGlnArgMetArgGlyThrPheValAlaProLeuPr	421	Db		540	540
QY	GATCCATCTCGGAACTCTAGACGCTTGTGCTCGCAGCGGCTCTGGAGCAAACT	1322		422	QY	TTGTTAGACGACGAGCGAGTCCCTTAGAAGAAGTCCCTCGCTCGCTCGCAGCGAAGCT	2402	
Db		421	oIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe	441	Db		540	540
QY	TATCGGAACCGCAACTCTGTTGCTCTCTCGAAATACACCTCTCTTCCATGGCTGT	1382		442	QY	CTCAATCGCGCGTTCGAGAAGATCTCAATCTCGGGAATCTCAACCTTAGTATTCCTTGG	2462	
Db		441	uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTyrLeuLe	461	Db		540	540
QY	AGGTTGTCTGCAACTGGAATCTCGCGGAGAGCTCTTGTGCTAGCTCCGCTCGCGCT	1442		462	QY	ACTCATAGGTGGGAAAATTTACTGGGCTTTATCTTCTACTGTACTCTCTTTAATCCC	2522	
Db		461	uGlyCysAlaAlaAsnTyrIleLeuArgGlyThrSerPheValTyrValProSerAlaLe	481	Db		540	540
QY	GAATCCCGGAGCAGCCCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTTCTTCATCT	1502		482	QY	GAGTGGCAATTCCTTCTCTCTTCTCATTTCATTAACAAGAGACATTTAATAGATGT	2582	
		1443		483	Db		540	540
				484	Db		540	540
				485	Db		540	540
				486	Db		540	540
				487	Db		540	540
				488	Db		540	540
				489	Db		540	540
				490	Db		540	540
				491	Db		540	540
				492	Db		540	540
				493	Db		540	540
				494	Db		540	540
				495	Db		540	540
				496	Db		540	540
				497	Db		540	540
				498	Db		540	540
				499	Db		540	540
				500	Db		540	540
				501	Db		540	540
				502	Db		540	540
				503	Db		540	540
				504	Db		540	540
				505	Db		540	540
				506	Db		540	540
				507	Db		540	540
				508	Db		540	540
				509	Db		540	540
				510	Db		540	540
				511	Db		540	540
				512	Db		540	540
				513	Db		540	540
				514	Db		540	540
				515	Db		540	540
				516	Db		540	540
				517	Db		540	540
				518	Db		540	540
				519	Db		540	540
				520	Db		540	540
				521	Db		540	540
				522	Db		540	540
				523	Db		540	540
				524	Db		540	540
				525	Db		540	540
				526	Db		540	540
				527	Db		540	540
				528	Db		540	540
				529	Db		540	540
				530	Db		540	540
				531	Db		540	540
				532	Db		540	540
				533	Db		540	540
				534	Db		540	540
				535	Db		540	540
				536	Db		540	540
				537	Db		540	540
				538	Db		540	540
				539	Db		540	540
				540	Db		540	540

Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGCAATATCTTACTTAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGlnHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTGAAAGATATGTCAAAGAATTGTGGGTCTTTTGGGCTTTCG 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGGCCCTTTTACCAATGCGCTATCTCGCTTGATGCTTATATGATGATGATATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCACCTTTCTCGCAACTTACAGGCTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGCTGAGCGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGATCGGGCTTGGCCATAGGCCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCT 1259
Db 703 oThrGlyTTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGGATCCATCTGCGGAACCTCTAGCAGCTGTGTTTGTCTCGCAGCGGTCTGGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGCAACTCTGTGTTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGGTGTGTCGCAACTCGATCTCGCGCGGACGTCTCTTGCTACGTCCCGFCGCG 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGAGCAGCCGCTCTCGGGCGGTGTTGGGGCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCGGTTCGCGCCACACACCGCGGCACCTCTCTTTACGCGGTCTCCCGCTATGTGCG 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACTTCGACGTGCGCATGAGACCAACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843

RESULT 5
Q4FD83_HPBVO PRELIMINARY; PRT; 843 AA.
ID Q4FD83; AC Q4FD83;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HK343;
RX PubMed=15897987; DOI=10.1086/430324;
RA Chan H.L., Tsui S.K., Tee C.H., Ng E.Y., Au T.C., Yuen L.,

RA Bartholomeusz A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C";
RL J. Infect. Dis. 191:2022-2032(2005).
DR EMBL; DQ089793; AAZ05280.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94494 MW; BB50233DFEBA4E29 CRC64;
Alignment Scores:
Pred. No.: 8 92e-202 Length: 843
Score: 2822.00 Matches: 530
Percent Similarity: 98.5% Conservative: 3
Best Local Similarity: 97.9% Mismatches: 7
Query Match: 47.7% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x Q4FD83_HPBVO (1-843)
QY 1 CTCACAACATTCACCAAGCTCTGCTAGATCCCAAGGTGAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCCGACTACTGCTCTCCATATGCTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGGACCTCGCACCAACATGAGAACACAAACATCAGGATTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTGTTACAGGGGGGTTTTCTCGTTGACAAAGATCTCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGAGTCTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGAGGACCCACGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGGCAAAATTCGAGTCCCCCACTCCAATCACTCAACCACTCTTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGCGTATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTTGTGTTCTTCGAGTACCAAGGTATGTTGGCCCGTTCCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGCGGGCCATGCAAGACCTGCAAGCTGCACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTCGGACCGGAAATCTCACATT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTCGGGCTTCGCGAAGATTCTTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCGGCTCAGTTTACTAGTGCATTTGTTTTCAGTGGTTCCTAGGGCTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGGAAAGTCTGTACAACTCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATATACAAATTTCTTTTGTCTTTGGGTATACATTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583

QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACATCTCATGGATATGTAATGGAAGTTG 899
Db |||||
QY 583 oasnlvsrthrsyargtrpGlytyrSerLeuasnPhemetGlytyrValleGlysertr 603
Db |||||
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTTCGAAAACGTCC 959
Db |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheargLysLeuPr 623
QY 960 TGTAAATAGACTTATGATGTGGAAGATATGTCAAGAAATGTGGGTCTTTGGCTTTGC 1019
Db |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAAATGTGCTATCTGCTGCTGCTGCTTATATCATGTATACATC 1079
Db |||||
QY 643 aAlaProPheThrGlnCysGlytyrProAlaLeuMetProleuTyraLacysIleGlnSe 663
QY 1080 TAACGAGCTTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAAATATCTCAA 1139
Db |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrrLysAlaPheLeuCylysgLntyrLeuAs 683
QY 1140 CTTTATACCCCTTGGCCGGCAACGGTCCGGTCTCTGCCAAGTGTGTGTGACGCAACCCC 1199
Db |||||
QY 683 nLeuTyrrProValAlaargGlnArgserGlyLeuCysGlnValPheAlaaspAlaThrPr 703
QY 1200 CACTGGATGGGCTTGGCCATAGCCATCAGCGCATGCTGGAACTTTCTGGCTCTCT 1259
Db |||||
QY 703 oThrGlytrpGlyLeuAlaIleGlyHisGlnArgMetargGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATACCTCGGGAACCTCTAGCAGCTGTGTTTCTCGCAGCCGGTCTGGAGCAAA 1319
Db |||||
QY 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaargSerArgserGlyAlaLy 743
QY 1320 ACTTATCGGAACCGAACAACCTGTGTGCTCTCTCGGAATPACACCTCTCTTCCATGGCT 1379
Db |||||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrrThrSerPheProTrpLe 763
QY 1380 GCTAGGGTGTCTGCCAAGTGTGCTGGGGGACGTCCTTGTCTACGTCCTCGCTGGC 1439
Db |||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValtyrValProSerAl 783
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGGCGGTGTTGGGGCTCTACCGTCTCCCTTTCTCA 1499
Db |||||
QY 783 aLeuAsnProAlaaspAspProSerArgGlyArgLeuGlyLeuTyrrArgProLeuLeuHi 803
QY 1500 TCTGCGTTCGGCGGACACCGGGCGCACTCTCTTTACGCGTCTCCCGTANGTCC 1559
Db |||||
QY 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGTGTGCACCTTGGCTTGCACCTCTGCACCTGCGATGCGAGACCAC 1619
Db |||||
QY 823 oSerHisLeuProaspArgValHisPheAlaSerProleuHisValAlaTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843

RESULT 6

ID Q69026 HPBV0 PRELIMINARY; PRT; 540 AA.
AC Q69026
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P gene product (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVT;
RX MEDLINE=87059755; PubMed=3783127;
RA Okamoto H., Imai M., Shimozaki M., Hoshi Y., Iizuka H., Gotanda T.,

RA Tauda F., Miyakawa Y., Mayumi M.;
RT "Nucleotide Sequence of a Cloned Hepatitis B Virus Genome, Subtype
RT ayt:Comparison with Genomes of the Other Three Subtypes.";
RL J. Gen. Virol. 67:2305-2314(1986).
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL; X04615; CAA28286.1; -; Genomic_DNA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0003987; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000477; RVase.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00078; RVT_1; 1.
DR PRODOM; PD00814; DNAPol_viral_C; 1.
DR PROSITE; PS00878; RT_POL; 1.
FT NON TER 1
SQ SEQUENCE 540 AA; 60365 MW; 80446CA7C80F998A CRC64;

Alignment Scores:

Pred. No.: 6.02e-201 Length: 540
Score: 2810.00 Matches: 527
Percent Similarity: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.57% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q69026_HPBV0 (1-540)

QY 1 CTCACACACATTTCCACCAAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db |||||
QY 1 LeuHisAsnIleProSerSerAlaArgSerGlnSerGlyProIlePheSerCys 20
QY 61 TGTGGCTCAGTTCGGAAACAGTAAACCTGTTCCGACTACTCGCTCTCCCATATCGTC 120
Db |||||
QY 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrrCysLeuThrHisIleVal 40
QY 121 AATCTTCTCGAGACTGGGACCCCTGCACCGAATCGAGAACACAAATCATGAGATTCCT 180
Db |||||
QY 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGACCCCTCTCGTGTATACAGCGGGGTTTTCTCGTTGCAAGAATCTCTCAATACC 240
Db |||||
QY 61 ArgThrProAlaargValThrGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGCACCACGTCTC 299
Db |||||
QY 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
QY 300 CTGGCCAAAATTCGAGTCCCAACCTCCCAATCACTCACCAACCTCTGTCTCTCAATTT 359
Db |||||
QY 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTGCTATCGTGGATGTGTCTGGCGGTTTTTATCATATTCCTCTTCTATCTCTCTGC 419
Db |||||
QY 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrrHisIleProleuHisProAlaAl 140
QY 420 TATGCTCATCTCTTCTGTCTCTCTCTGACTACCAAGGTATGTTCCCGTTTGTCTCTC 479
Db |||||
QY 140 aMetProHisLeuLeuValGlySerGlyLeuProArgTyrrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGAACATCAACCAACGAGCAGCGGGGCGCATGCAAGACCTGCAGCATCTCTCTC 539
Db |||||
QY 160 rThrSerArgAsnIleAsnTyrrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
QY 540 AAGAAACTCTAGCTTTCCCTCTTGTGTGTGTATGATACAAACCTTCGGACGGAACTGCATT 599
Db |||||
QY 180 rArgAsnLeuTyrrValSerLeuLeuLeuTyrrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTCCATCCCATCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db |||||


```

Db      200 uTrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY      660 GTTTCCTCCGGCTCAGTTACTAGTCCCAATTTGTTTCAGTGGTTCTGAGGGCTTTCCCA 719
Db      220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValAlaArgAlaPheProH 240
QY      720 CTGTTGGCTTTTCAGTTATATGATGTGTGTTATTTGGGGGCGAAGTCTGTCAACATCT 779
Db      240 sCyLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHis 260
QY      780 TGAGTCCCTTTTACCTCTATACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
Db      260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY      840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTCATGGGATATGTAATTTGGAAGTTG 899
Db      280 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleLysTr 300
QY      900 GGTACTTTTACCGCAGGACATATTTGTTACTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
Db      300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
QY      960 TGTAAATAGACCTATTGATTGGAACTATGTCMAAGATTGCGGTCCTTTGGGCTTTGCC 1019
Db      320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY      1020 TGCCCTCTTTTACCAATGGCTATCTCCTCTGTATGCTTATATGTCATGTATATCAATC 1079
Db      340 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 360
QY      1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db      360 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLysLeuAs 380
QY      1140 CCTTTACCGCTTGGCCGCAACGTCGGTCTCTGCCAAGTGTCTGCTGAGCAACCC 1199
Db      380 nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY      1200 CACTGATGGGCTTCGGCATAGGCCATCAGCGCATGGCTGGAACCTTCTGCTCCTCT 1259
Db      400 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
QY      1260 GCGATCCATATCTCGGAATCTCTAGCAGCTTCTTTTGTCTCGCAGCGGTCTCGAGCAA 1319
Db      420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 440
QY      1320 ACTTATCGGAACCGCAACACTCTGTTGCTCTCTCGGAATATACCTCTCTTCCATGGCT 1379
Db      440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 460
QY      1380 GCTAGGGTGTGCTGCGCAACTGGATCTCGCGCGGAGCTCTCTTGTCTAGCTCCGTCGGC 1439
Db      460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 480
QY      1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCGCGTCTTGGGGCTCTACCGTCCCTCTTCA 1499
Db      480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHi 500
QY      1500 TCTGCGGTTCCGGCGGACGACGCGGCGACCTCTCTTTAGCGGCTCTCCCGTATGTGCC 1559
Db      500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 520
QY      1560 TTTCTATCTCGCGACCGTGTGACCTTCTGCTTCACTCTGACCTCTGACGTGCGAGAGACC 1619
Db      520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY      1620 G 1620
Db      540 O 540

```

RESULT 7
Q9QN49_HPBVO

ID Q9QN49_HPBVO PRELIMINARY; PRT; 843 AA.

```

AC      Q9QN49;
DT      01-MAY-2000 (TRENBLrel. 13, Created)
DT      01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT      01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE      Polymerase.
OS      Hepatitis B virus.
OC      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
ON      NCBI_TaxID=10407;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=genotype C;
RX      MEDLINE=20528425; PubMed=11074476;
RX      DOI=10.1002/1096-9071(200012)62:4<471::AID-JMV12>3.3.CO;2-O;
RA      ROKuhara A., Tanaka B., Yagi S., Mizokami M., Hashikura Y.,
RA      Kawasaki S., Kiyosawa K.;
RT      "De novo infection of hepatitis B virus in patients with orthotopic
RT      liver transplantation: analysis by determining complete sequence of
RT      the genome.";
RL      J. Med. Virol. 62:471-478(2000).
DR      EMBL; AB026815; BAA85065.1; -; Genomic_DNA.
DR      GO; GO:0003677; P:DNA binding; IEA.
DR      GO; GO:0003887; P:DNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0004519; P:endonuclease activity; IEA.
DR      GO; GO:0016787; P:hydrolase activity; IEA.
DR      GO; GO:0004523; P:ribonuclease H activity; IEA.
DR      GO; GO:0003723; P:RNA binding; IEA.
DR      GO; GO:0003964; P:RNA-directed DNA polymerase activity; IEA.
DR      GO; GO:0016740; P:transferase activity; IEA.
DR      GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR      InterPro; IPR001462; DNAPol_viral_C.
DR      InterPro; IPR000201; DNAPol_viral_N.
DR      InterPro; IPR000477; RVTse.
DR      Pfam; PF00336; DNA_pol_viral_C; 1.
DR      Pfam; PF00242; DNA_pol_viral_N; 1.
DR      Pfam; PF00078; RVT_1; 2.
DR      ProDom; PD000814; DNAPol_viral_C; 1.
SQ      SEQUENCE 843 AA; 94462 MW; 9D741241062DEAAE CRC64;

```

Alignment Scores:

Pred. No.:	9,92e-201	Length:	843
Score:	2808.00	Matches:	527
Percent Similarity:	98.52%	Conservative:	6
Best Local Similarity:	97.41%	Mismatches:	7
Query Match:	47.54%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x Q9QN49_HPBVO (1-843)

```

QY      1 CTCCACCAATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
      |||:::|||||
Db      304 LeuHisHisIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY      61 TGTGTGCTCAGTTCGGGAACAGTAAACCTGTCGACTACTGCTCTCCCATATCGTC 120
      |||:::|||||
Db      324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuThrHisIleVal 343
QY      121 AATCTTCTCAGGACTGGGGACCTCGACCGAATCGGAGAACACAAACATCAGGATTCCT 180
      |||:::|||||
Db      344 AsnLeuLeuGluAsnTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY      181 AGGACCCCTCGTCTGTTTACAGCGGGGTTTTTCTCGTTTGACAAGAAATCTCAATACC 240
      |||:::|||||
Db      364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY      241 GCAGAGCTAGACTCG-GTGAGCTTCTCTCAATTTCTAGGGGGAGCACCCACGTTGTC 299
      |||:::|||||
Db      384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY      300 CTGGCCAAAATTCGACAGTCCCAACCTCCAATCAGTCACCAACCTCTGTCTCTCCCAATTT 359
      |||:::|||||
Db      403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY      360 GTCCTGGCTATCGCTGGATGTGTCTCGGGGTTTTTATCATATTTCTCTTCTATCTGCTGC 419

```



```
Db 423 uSerfPrLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGGTCTTCTGGAATACCAAGGTATGTTGCCGTTTGTCTTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCAGGAACATCAACACAGCAGCGGGCCATCAGACCTGACGACTGCCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTACGCTTCTTGTGGTGTACAAAACCTTCGACGCGAAACTGCACATT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCATCCCATCATCTCGGGCTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTGGCTCAGTTACTAGTGCATTTGTTCACTGCTTCGTAGGCTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTCGGGCGAAGTCTGTACAACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTCTTCTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAACCAAGCTTGGGCTACTCCTTAACCTCATGCGATATGTAATTTGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTTACCGCAGCAACATATGTTACTAAACTCAAGCAATGTTTTCGAAACCTGC 959
Db 603 pGlyThrLeuProGlnGlnHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGGCTATCTGCTGATGCTTATATGATGATGATGATGATGAT 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
Qy 1140 CTTTACCCGTTGCGCGCAACCGTCCGGTCTCTGCAAGTGTGTTGCTGACGCAACCCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGATGGGCTTGGCCATAGCCATCAGCGATGCGTGGAAACCTTTCTGGCTCTCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTCGGCACTCTTAGCAGCTTGTGTTGCTCGCAGCGGCTCTCGAGCAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGACCACTCTGTGTCTCTCTCGGAAATACACCTCTCTTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
Qy 1380 GCTAGGTGTGCTGCCAACTGGATCTGCGCGGAGCGTCTTGTCTACGTCGCGTCCGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCGCGAGCAGCCGCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTCTTCA 1499
```

```
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
Qy 1500 TCTGCCGTTCCGCCCGCACACCGGCGCACTCTCTTTACGGGCTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTTCTCATCTGCCGCGCGTGTGCACCTTCGCTTCCACCTCTGCACGTGCGATGGAGACACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 O 843

RESULT 8
DPOL_HPBVR STANDARD; PRT; 843 AA.
ID DPOL_HPBVR AC P03157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 13-SEP-2005 (Rel. 48, Last annotation update)
DE P protein [includes: DNA-directed DNA polymerase (EC 2.7.7.7); RNA-
directed DNA polymerase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN Names:P;
OS Hepatitis B virus (subtype adr).
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RX NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA MEDLINE=83158919; PubMed=6300776;
Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adv.";
RL Nucleic Acids Res. 11:1747-1757(1983).
CC -1- CATALYTIC ACTIVITY: Deoxynucleoside triphosphate + DNA(n) =
diphosphate + DNA(n+1).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphomonoester.
CC -1- SIMILARITY: Contains 1 reverse transcriptase domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
EMBL: V00867; -; NOT_ANNOTATED_CDS; Genomic_DNA.
PIR: A00704; JDVLYR.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR00201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C; 1.
DR Pfam; PF00242; DNA_pol_viral_N; 1.
DR Pfam; PF00078; RVT_1; 1.
DR ProDom; PD000814; DNAPol_viral_C; 1.
DR PROSITE; PS50878; RT_POL; 1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase;
Endonuclease; Hydrolase; Multifunctional enzyme; Nuclease;
Nucleotidyltransferase; RNA-directed DNA polymerase; Transferase.
FT DOMAIN 357 600 Reverse transcriptase.
SQ SEQUENCE 843 AA; 94400 MW; A6B2D490839C48BB CRC64;

Alignment Scores:
Pred. No.: 1,18e-200 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x DPOL_HPBVR (1-843)
```

QY	4	CACAACTATCCACCAAGCTCTGCTAGATCCAGGGGTGAGGGGCTATATATTTTCTCGCTGG	63
Db	305	HIAAAILeProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp	324
QY	64	TGCTCCAGTTCGGGACAGTAACCTGTTCGAGCTACTCTCTCCCATATCGTCAAT	123
Db	325	TrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuThrHisIleValAa	344
QY	124	CTTCTCAGGACTGGGACCCCTGCACCGAATATGGAGAACACAACTACGATTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg	364
QY	184	ACCCCTGCTCGTGTATTACGCGGGGGTTTTCTCGTTGACAAGAATCTCAATAACGCA	243
Db	365	ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAenProHisAenThrThr	384
QY	244	GAGTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTTCTTG	302
Db	385	GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal- SerTr	404
QY	303	GCCAAATTCGCTAGTCCCAACTCCAACTCACCTCACCAACTCTGTGCTCCCAATTTGTC	362
Db	404	pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe	424
QY	363	CTGGCTATCGCTGGATGTCTCGGGGGTTTTTATCATATTCTCTTCATCTCTGCTGAT	422
Db	424	rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe	444
QY	423	GCCTCATCTTCTGTGTGGTCTTCTTGACATACCAAGGTATGTGGCGTTGTCTCTCTAC	482
Db	444	tProHisLeuLeuValGlySerGlyLeuProArgTyrValAlaArgLeuSerSerTh	464
QY	483	TTCCAGGAACATCAACACACAGCAGCGGGCCATCGAAGACTCTGCAGACTCTCTGCTCAAG	542
Db	464	rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr	484
QY	543	GAAACTCTAGCTTTCCCTCTTGTGTGTGTACAAAACCTTCGAGCGGAAACTGCACCTGTGA	602
Db	484	gAenLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyr	504
QY	603	TTCCCATCCCATCATCTCTGGGCTTCGCAAGATTCCTATGCGAGTGGGCTCAGTCGGTT	662
Db	504	rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh	524
QY	663	TCCTCGCTCAGTTTACTAGTGCATTTCTCAGTGTGTTCTGATGGCTTTCCCCCACTG	722
Db	524	eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy	544
QY	723	TTTGGCTTTTCAGTTATATGATGATGTGGTATTGGGGCGGAAGTCTGTACAACTTTGA	782
Db	544	sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl	564
QY	783	GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCCATA	842
Db	564	uSerLeuPheThrSerIleThrAenPheLeuLeuSerLeuGlyIleHisLeuAenProAs	584
QY	843	TAAACCAACAGTTGGGCTACTCCCTTAATTCATGGGATATGTAATTCGAAGTTGGGG	902
Db	584	nLysThrLysArgTrpGlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTrpGl	604
QY	903	TACTTTACCGCAGAAACATATTGACTTAAAACTCAAGCAATGTTTTTCGAAAACTGCTGT	962
Db	604	yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa	624
QY	963	AAATPAGACCTATTGATTGGAAAGATGTCAAGAAGATTTGTGGTCTTTTGGGCTTTGCTGC	1022
Db	624	iAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl	644
QY	1023	CCCTTTTACAAATGTGGCTATCTCGCTTGATGGCTTTATATGCATGTATACATCTAA	1082
Db	644	aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy	664
QY	1083	CGAGGCTTTTCACTTTCTCGCCAACTTACAAAGGCCCTTTCTGTGTGTAAAACAATATCTGAACCT	1142

Db	664	sgAlaPheThrPheSerProThrTyrLySAlaPheLeuCysLysGlnTyrLeuAsnLe	684
Qy	1143	TTATCCCGTGGCCGGCAACGGTCTCTGCAGTGTGTTGCTGACGCAACCCCCAC	1202
Db	684	uTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh	704
Qy	1203	TGGATGGGGCTTGGCCATAGCCATCAGCCATCAGCGCATGGCTGGGAACCTTTCTGGCTCTCTG	1262
Db	704	RGlyTriPGlyLeuAlaIleGlyHisArgAGMetArgGlyThrPheValAlaProLeuPr	724
Qy	1263	GATCCATACATGCGGAACCTCTACGAGCTTGTTTGTCTCGCAGCGGCTCTGGAGCAAACT	1322
Db	724	oIleHisThrAlaGluLeuAlaCysPheAlaArgSerArgSerGlyAlaLysLe	744
Qy	1323	TATCGGAACCGCAACACTCTGTCTCTCGAATAACACTCCTCTTCATCGGCTGCT	1382
Db	744	uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe	764
Qy	1383	AGGCTGTGCTGCCAACTGGATCTCTGGCGGGAGCTCTCTTGTCTACGTCCTCGCGCGCT	1442
Db	764	uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe	784
Qy	1443	GAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCACT	1502
Db	784	uAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLe	804
Qy	1503	GCGCTTCGGCGCCGACACGGGGCGCACTCTCTTACGGCGGTCTCCCGCTATGCGCTTC	1562
Db	804	uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe	824
Qy	1563	TCATCTCCCGACCGTGTGCACCTTCGCTTCACCTCTCAGCTCGCATGGAGACACCG	1620
Db	824	rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro	843

RESULT 9

Q97975 HPBV0 PRELIMINARY; PRT; 540 AA.

AC Q97975; DT 01-FEB-1997 (TrEMBLrel. 02, Created) DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update) DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update) DE ORF (Fragment). OS Hepatitis B virus. OC Viruses; Retro-transcribing viruses; Hepadnaviridae; OC Orthohepadnavirus. OX NCBI_TaxID=10407; [1] NUCLEOTIDE SEQUENCE. RX MEDLINE=95294549; PubMed=7775946; RA Uchida T., Gotoh K., Shikata T.; RT "Complete nucleotide sequences and the characteristics of two RT hepaditis B virus mutants causing serologically negative acute or chronic hepatitis B."; J. Med. Virol. 45:247-252(1995). RC -I- SIMILARITY: Contains 1 reverse transcriptase domain. DR EMBL; D16666; BAA04073.1; -; Genomic_DNA. DR GO; GO:0003677; F:DNA binding; IEA. DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA. DR GO; GO:0004523; F:ribonuclease H activity; IEA. DR GO; GO:0003723; F:RNA binding; IEA. DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA. DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA. DR InterPro; IPR001462; DNAPol_viral_C. DR InterPro; IPR000477; RVTse. DR Pfam; PF00336; DNA_pol_viral_C; 1. DR Pfam; PF00078; RVT_1; 1. DR ProDom; PD000814; DNAPol_viral_C; 1. DR PROSITE; PS50878; RT_POL; 1. FT NON_TER 1 SQ SEQUENCE 540 AA; 60346 MW; 7BBE285359DD39C5 CRC64;

Alignment Scores:

Pred. No.: 1.2e-200 Length: 540
Score: 2806.00 Matches: 526
Percent Similarity: 98.34% Conservative: 6
Best Local Similarity: 97.23% Mismatches: 8
Query Match: 47.50% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q97975_HPBVO (1-540)

QY 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTCCAGGGCTATATTTCTCTGC 60
DB 1 LeuHisAsnLeuProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGTGGCTCCAGTTCGGAACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
DB 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
QY 121 AATCTTCTCAGGACTGGGACCTCGACCGAATCGGAGAACACACATCAGGATTCCT 180
DB 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCCCTCTCGTGTACAGCGGGGTTTTCTCGTTGACAGAACTCCTCAATAACC 240
DB 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCGAGTCTAGACTCTG-GTGGACTTCTCAATTTTCTAGGGGAGCACCCACGTTTC 299
DB 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisIleVal-Se 100
QY 300 CTGGCCAAATTCGAGTCCCACTCCCACTCACTCACTCACTCACTCTGTCCTCAATTT 359
DB 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTGCTGCTATCGTGTGATGCTGGGGGTTTTATCATATTCCTCTTCTCATCTGCTGC 419
DB 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
QY 420 TATGCTCATCTTCTGTTGTTCTTCTGACTACCAAGTATGTTGCCGTTGCTCTC 479
DB 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGACATCAACACGACGCGGGCCATCGACCTGACAGCTGACGACTCTGCTC 539
DB 160 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 180
QY 540 AAGGAACTCTACGTTTCCCTCTGTTGCTGTACAAACCTTCGACGCGAACTGCACTT 599
DB 180 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTCCTCATCTCTGCTGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
DB 200 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTCTCTGCTCAGTTACTAGTGCATTTGTTGAGTGTTCAGTGTTCGTTAGGCTTCCCCCA 719
DB 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 779
DB 240 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTATTACCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 839
DB 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY 840 TAATAAACCAAGCTTGGGCTACTCCCTTAACCTTATCGGATATGATTAATGGAAGTTG 899
DB 280 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGGTACTTTTACCGGAGGAACATATGTACTAAAACTCAAGCAATGTTTTCGAAACCTGCC 959
DB 300 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 320

QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTGCGGTCTTTTGGGCTTTC 1019
DB 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGGCCCTTTTACACAATGTTGGCTATCTGCTTGTATGCTTATATGTCATGTATACAATC 1079
DB 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
QY 1080 TAAGCAGGCTTTCACCTTCTCGCAACTTCAAGGCTTCTGCTGTAAACAATATCTGAA 1139
DB 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 380
QY 1140 CCTTTACCCGTTTGGCCGCAAGGTCGCTCTCTCCCAAGTGTGTTGCTGACCGAACCCC 1199
DB 380 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
QY 1200 CACTGATGCGGCTTGGCCATAGGCCATCAGCCATGCGTGGAACTTCTTCTGCTCTCTCT 1259
DB 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
QY 1260 GCGATCCATCTACTGCGAACTCTCTAGCAGCTTGTGCTCGCAGCGGCTCTGGAGCAA 1319
DB 420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
QY 1320 ACTTATCGGAACGCAACTCTGTTCTCTCTCTCGAAATACACCTCTTCCATGGCT 1379
DB 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 460
QY 1380 GCTAGGCTGTGTCGCAACTGGAATCTGCGCGGAGCTCTTGTCTTACGCTCCGCTCGGC 1439
DB 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA 1499
DB 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGCTTCCGCGGCGACCGGCGGCGACCTCTCTTACGCGGCTCTCCCGGTATGTC 1559
DB 500 sLeuProPheArgProThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTTCTATCTGCGGACCGCTGTGCACTTCTGCTTCCCTCTGACGCTGCGATGGAGACCC 1619
DB 520 oSerHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
DB 540 O 540

RESULT 10
Q5KR39_HPBVO PRELIMINARY; PRT; 843 AA.
ID Q5KR39_HPBVO PRELIMINARY; AC Q5KR39;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
ON NCBI_TaxID=10407;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RP Sun X., Rokuhara A., Tanaka E., Gad A., Mutou H., Matsumoto A.,
RA Yoshizawa K., Kiyosawa K.;
RT "Nucleotide Mutations Associated With Hepatitis B e Antigen
RL J. Med. Virol. 76:170-175 (2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Rokuhara A.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB198077; BAD86600.1; -; Genomic DNA.

DR GO: GO:0003677; F:DNA binding; IEA.
 DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0004519; F:endonuclease activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0004523; F:ribonuclease H activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro: IPR001462; DNAPol_viral_C.
 DR InterPro: IPR000201; DNAPol_viral_N.
 DR InterPro: IPR000477; RVase.
 DR Pfam: PF00336; DNA_pol_viral_C; 1.
 DR Pfam: PF00342; DNA_pol_viral_N; 1.
 DR Pfam: PF00078; RVF1; 1.
 DR ProDom: PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 843 AA; 94446 MW; 80CD75F48581AC2F CRC64;

Alignment Scores:

Pred. No.: 1.4e-200 Length: 843
 Score: 2806.00 Matches: 526
 Percent Similarity: 98.15% Conservative: 5
 Best Local Similarity: 97.23% Mismatches: 9
 Query Match: 47.50% Indels: 2
 DB: Gaps: 0

US-10-761-006A-1 (1-3215) x Q5KR39_HPBVO (1-843)

QY 1 CTCCACAACTTCCACAGCTCTGTAGATCCAGGGTGAGGGCGCTATATTTTCCTGC 60
 DB 304 LeuHisAsnIleProProSerAlaArgProGlnSerGluGlyProIleLeuSerCys 323
 QY 61 TGTGGCTCCAGTTCGGGAACAGTAACCTCTGTCGACTACTGCTCTCCCATATCGTC 120
 DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
 QY 121 AATCTTCTGAGGACTGGGGACCTGACCGAACATATGAGAGAACACAACTCAGGATTCCT 180
 DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
 QY 181 AGGACCCCTGCTGTTACAGCGGGTTCCTGCTTGACAGAACTCCTCACATACC 240
 DB 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
 QY 241 CGAGAGTCTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGAGACCCACGCTGTC 299
 DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
 QY 300 CTGGCCAAAATTGCGAGTCCCAACCTCAATCCTCACCACCTCTGCTCTCCAAATT 359
 DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
 QY 360 GTCTGGCTATCGCTGGAGTGTCTGCGGGTTCATATATTCCTTTCATCTGCTGC 419
 DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
 QY 420 TATGCTCATCTTCTTGTGTTCTTCTGACTTACCAAGGTATGTTGCCGTTGTCCTC 479
 DB 443 aMetProHisLeuLeuValValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
 QY 480 TACTTCCAGGAACATCAACACAGCAGCGGGGCATGCAAGACCTGACGACTCTCTGCTC 539
 DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
 QY 540 AAGGAACCTAGTTCCTCTGTTGCTGTGTACAAAACCTTCGGACGGAACTGCACATT 599
 DB 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle 503
 QY 600 GTATTCCTCCATCCATCATCTCTGGGCTTTCGCAAGATTCTATGGAGTGGGGCTCAGTCC 659
 DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 QY 660 GTTCTCTCGGCTCAGTTTACTAGTCCCAATTGTTTCAGTGGTTCGTAGGGCTTTCCCCCA 719

DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
 QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGGCGAAGTCTGTACAACATCT 779
 DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisIle 563
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTAAACCC 839
 DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 QY 840 TAAATAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATGTAAATGGAGTTG 899
 DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
 QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
 DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAATTGTGGGTCTTTTGGGCTTTC 1019
 DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TGCCCTTTTACACAAATGTCCTATCTGCTGCTGCTGCTTATATGATGATGATACATC 1079
 DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
 QY 1080 TAAGCAGGCTTTTCACTTCTCGCCAACTTCAAGGCTTCTGTGTAACAATATCTGAA 1139
 DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
 QY 1140 CCTTTACCCGTTGCCCGCAACGTCGCGTCTCTGCCAAGTGTGTGTCGACCAACCC 1199
 DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGGATGGGGCTGGCCATAGGCCATCAGCGCATGGCTGCAACCTTCTCGCTCCTCT 1259
 DB 703 oThrGlyTrpGlyLeuAlaValGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCGGATCCATCTAGCGGAACCTCTAGCAGCTGTTTGTCTCGCAGCCGGTCTGGAGCAA 1319
 DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaL 743
 QY 1320 ACTTATCGGAACCGCAACTCTGTTGTCCTCTCTCGGAATAACACTCTCTTCCATGGCT 1379
 DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
 QY 1380 GCTAGGTGTGTGTCCAACTGATCTCGCGGGAGCTCTTGTGTCTAGCTCCGCTCGGC 1439
 DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTTCTTCA 1499
 DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
 QY 1500 TCTGCGGTTCCGGCCGACACCGGGCGCACCTCTCTTTACGGGTCTCCCGGTATGTGCC 1559
 DB 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGGTGTGCACCTTCGCTTCACCTCTGCGACGTGCGATGGAGACC 1619
 DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 DB 843 o 843
 RESULT 11
 Q7TDR3 HPBV0
 ID Q7TDR3 HPBV0 PRELIMINARY; PRT; 843 AA.
 AC Q7TDR3-
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

[illegible]

```

Db      763 uLeuGlyCyAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
Qy      1440 GCTGAATCCCGCGGACACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCTCA 1499
Db      783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHi 803
Qy      1500 TCTGCCGTTCCGGCCGACACGCGGCGACCTCTCTTTAGCGGTCTCCCGGTATGTGCC 1559
Db      803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
Qy      1560 TTCTCACTCTCGGACCGTGTGCACTTCTGCTTCACTCTGCACTGTCATGGAGACCAACC 1619
Db      823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy      1620 G 1620
Db      843 o 843

RESULT 13
Q4FD65_HPBVO
ID Q4FD65_HPBVO PRELIMINARY; PRT; 843 AA.
AC Q4FD65;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polymerase.
GN Name=P;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RC NUCLEOTIDE SEQUENCE.
RA CHAN H.K., TSE C.H., NG E.Y., AU T.C., YUEN L.,
RA Bartholomew A., Leung K.S., Lee K.H., Locarnini S., Sung J.J.;
RT "Epidemiological and virological characteristics of 2 subgroups of
RT hepatitis B virus genotype C.";
RL J. Infect. Dis. 191:2022-2032 (2005).
DR EMBL; DQ089798; AN205298.1; -; Genomic DNA.
SQ SEQUENCE 843 AA; 94318 MW; 0C749903F281504A CRC64;

Alignment Scores:
Pred. No.: 2.35e-200 Length: 843
Score: 2803.00 Matches: 525
Percent Similarity: 98.34% Conservativity: 7
Best Local Similarity: 97.04% Mismatches: 8
Query Match: 47.45% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q4FD65_HPBVO (1-843)
Qy 1 CTCACAAACATCCACCAAGCTCTGTAGATCCAGGGTGGGGCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
Qy 61 TGGTGGCTCCAGTTCGCGACAGTAACCCGTTCGACACTCCCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTTCGACCCGAAATGAGAACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTCGTGTATACAGGCGGGGTTTTCTCGTTGACAGAAATCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGGAGCACCCAGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403

```

```

Qy 300 CTGGCCAAAATTCCGAGTCTCCCAACCTCAATCAGTCAACCACTCTGTCTCTCCAAATT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTCTGGCTATCGCTGGATGTGTCTCGCGGCTTTTATCATATTCCTCTTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGTGTCTTCTGGAGTACCAAGGTATGTTCGCCGTTGTGCTCT 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProGlyTyValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCAAGACTGCACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTrpGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTACGTTTCCCTCTGTGTGTGTACAAAACCTTCGAGCGGAAACTGCACATT 599
Db 483 rArgAsnLeuTyValSerLeuLeuLeuLeuTyLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCTCCATCATCTCTGGGCTTTCGCAAGATTCTCTATGGAGTGGGCTCTAGTCC 659
Db 503 uTySerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCTGGCTCAGTCTTACTAGTGCATTTCTTTCAGTGTGCTGAGGCTTCCCTCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTATATGATGATGTGGTATTGGGGCGCAAGTCTGTACAACATCT 779
Db 543 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTCTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuTyThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACAAACGTTGGGCTACTCCCTTAACCTTCGAGATATGATTAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 603
Qy 900 GGTACTTTTACCCGACGACATATGCTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACCTATTGATTGGAAGTATGCTCAAGAATTTGTGGGTCTTTTGGCTTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGGCTATCTGCTGCTTGTATGCTTATATGATGATGATGATGATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
Qy 1080 TAGACGGCTTTCACCTTCTCGCAACTTACAGGCTTCTCTGTGTAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLeuAs 683
Qy 1140 CCTTTACCCCGTTCGCGCAACCGTCTCTCCAAAGTGTGTTGCTGACGCAACCC 1199
Db 683 nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGCTTGGCCATAGCCCATAGCCATGCTGGAACTTCTGCTGCTCTCTCT 1259
Db 703 oThrGlyTyPheLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTGCGGAACCTCTGACGCTTCTTTTGTCTGCGACCGGCTCTGAGCAAA 1319
Db 723 uProIleHisThrAlaGlnLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGCAACTCTGTGTGTCTCTCTCGGAAATATACACCTCTCTTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTyle 763

```


QY 1380 GCTAGGGTGTGTGTCACAACTGGATCCTGCGCGGAGCGTCTTTGTCTACGTCCCGTCGGC 1439
 Db 763 uleuGlyCyAlaAlaAAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
 QY 1440 GCTGAATCCCGGAGGAGCCGCTCTCGGGCGGTTGGGGCTCTACCGTCCCGCTTCTCA 1499
 Db 783 aleuAAsnProAlaAAspArgProSerArgGlyLeuGlyLeuPheArgProLeuLeuHi 803
 QY 1500 TCTGCGGTTCCCGGCGACACACCGGCGCACCTCTCTTTACGCGGTCTCCCGCTATGTGCC 1559
 Db 803 eLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGGTGTGCACCTTCTGCTTCACTCTGCACTGCGAGTGAAGACCCACC 1619
 Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTIPArgProPr 843
 QY 1620 G 1620
 Db 843 o 843
 RESULT 14
 Q81165 HPBV0 PRELIMINARY; PRT; 838 AA.
 AC Q81165;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE DNA polymerase.
 OS Hepatitis B virus.
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 OX NCBI_TaxId=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RX MEDLINE=93096607; PubMed=1461746;
 RA Mukaida M.;
 RT "The complete nucleotide sequence of hepatitis B virus, subtype adr
 (SRADR) and phylogenetic analysis.";
 RL Nucleic Acids Res. 20:6105-6105(1992).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=subtype adr;
 RA Uchida T., Aye T.T., Shimojima M., Gotoh K., Shikata T.;
 RT "Full-length nucleotide sequence of a hepatitis B virus (HBV) mutant
 isolated from a patient with acute hepatitis who did not exhibit
 serological markers for HBV infection.";
 RL Int. Hepatol. Commun. 2:70-73(1994).
 DR EMBL; D16685; BAA04072.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
 DR InterPro; IPR001462; DNAPol_viral_C.
 DR InterPro; IPR000201; DNAPol_viral_N.
 DR InterPro; IPR000477; RVase.
 DR Pfam; PF00336; DNA_pol_viral_C; 1.
 DR Pfam; PF00242; DNA_pol_viral_N; 1.
 DR Pfam; PF00078; RVT_1; 2.
 DR ProDom; PD000814; DNAPol_viral_C; 1.
 SQ SEQUENCE 838 AA; 93902 MW; A686F42940C3B806 CRC64;

Alignment Scores:

Pred. No.: 2,78e-200 Length: 838
 Score: 2802.00 Matches: 525
 Percent Similarity: 98.33% Conservative: 6
 Best Local Similarity: 97.22% Mismatches: 8
 Query Match: 47.44% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x Q81165_HPBVO (1-838)
 QY 4 CACAACTTCCACCAAGCTCTGTAGATCCACAGGTGAGGGCCCTATATTTTCTGCTGG 63
 Db 300 HisAnIlePProPProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 319
 QY 64 TGGCTCCAGTTCGGGAACAGTAAACCTGTTCGGACTACTGCTCTCCCATATCGTCAAT 123
 Db 320 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAAsn 339
 QY 124 CTTCTCGAGGACTGGGAGCCCTGACCGAACAATGAGAGACACAAATCAGGATTCCTAGG 183
 Db 340 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 359
 QY 184 ACCCTGCTGCTGTTTACAGGGGGGTTTTCTGTTGACAAAGTCTCACAAATCCTCACAAATCCGCA 243
 Db 360 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr 379
 QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGCACCCACGTTCTCTG 302
 Db 380 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 399
 QY 303 GCCAAAATTCGAGTCCCAACCTCCAATCACTACCAACCTCTTGTCTCCTCAATTTGTC 362
 Db 399 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 419
 QY 363 CTGGCTATCGCTGAGTGTCTCGCGGCTTTTATCATATTCCTCTTCTCTCTGCTGCTAT 422
 Db 419 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 439
 QY 423 GCCTCATCTCTTGTGTGTCTTCTGGACTACCAAGGTATGTTGCCGTTTCTCTCTCTAC 482
 Db 439 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 459
 QY 483 TTCCAGGAAACATCAACACACCGAGCGGCCCATGCAAGCTTCGACGACTCTCTGCTCAAG 542
 Db 459 rSerArgAsnIleAAsnTyrHisGlyThrMetGlnAspLeuHisAspSerCysSerAr 479
 QY 543 GAAACTCTAGCTTCCCTCTGCTGTACAAAACCTTCGACGCGAAATCGCACCTGTGA 602
 Db 479 GAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 499
 QY 603 TTCCCATCCCATCATCTGGGCTTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTT 662
 Db 499 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 519
 QY 663 TCTCTGGCTCAGTTTACTAGTCCCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCACATG 722
 Db 519 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCy 539
 QY 723 TTTGGCTTTTCAGTTATATGATGTGTGTATTGGGGGGAAGTCTGTACAACATCTTGA 782
 Db 539 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 559
 QY 783 GTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGGTATATACATTTAAACCTTAA 842
 Db 559 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 579
 QY 843 TAAACCAAAACGTTGGGCTACTCCCTTAACTTTCATGGATATGTAATTTGGAAGTTGGGG 902
 Db 579 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 599
 QY 903 TACTTTACGCGAGGACATATTGTACTAAACCTCAAGCAATGTTTTTCGAAAACTGCTGT 962
 Db 599 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 619
 QY 963 AAATAGACCTATTGATTGGAAAGTATGTCAAAAGATTTGGGGTCTTTTGGGCTTTTCTGTC 1022
 Db 619 lAsnArgProIleAAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 639
 QY 1023 CCCTTTTACAAATGTGGCTATCTCGCTTGTATGCTTTTATGATGATATATCAACTAA 1082

Db 639 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerIy 659
Qy 1083 GCAGGCTTTCACCTTCTCGCAACTTACAGGCTTCTGTGTAAACAATATCTGAACCT 1142
Db 659 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHisLe 679
Qy 1143 TTATCCCGTTCGCGGCAAGCTCCGCTCTCTCCAGTGTGTCTGAGCAACCCCGAC 1202
Db 679 uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProth 699
Qy 1203 TGGATGGGCTTGGCCATAGCCATAGCGATCGATCGCTGGAACTTTCTGGCTCTCTGCC 1262
Db 699 rGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuPr 719
Qy 1263 GATCCATATCGCGAACTCTAGCAGCTTGTGTCTCGAGCGGTCTGGAGAAAACCT 1322
Db 719 oIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 739
Qy 1323 TATCGGACGACAACTCTGTCTCTCTCTCGGAATACACCTCTCTCCATGGCTGCT 1382
Db 739 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 759
Qy 1383 AGGTGTGCTGCGCAACTCGATCTCTGCGCGGACGCTCTGTCTACGTCCTCGCGCGCT 1442
Db 759 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 779
Qy 1443 GAATCCCGCGGACGACCGCTCTCGGCGCGGTGTGGGCTCTACCGTCCCTTCTTCATCT 1502
Db 779 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLe 799
Qy 1503 GCGTTCGCGCGACGACGCGGCGACCTCTCTTACGCGGTCTCCCGGTATGCGCTTC 1562
Db 799 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 819
Qy 1563 TCATCTCCGCGACGCTGCACTTCGCTTCACTCTGACCTGTCAGCTGCGATGGAGACCGG 1620
Db 819 rHisLeuProSerArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 838

RESULT 15
O42041_HPBVO
ID O42041_HPBVO PRELIMINARY; PRT; 843 AA.
AC O42041; 1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DNA polymerase.
OS Hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
(1)
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=97160188; PubMed=9007704; DOI=10.1016/S0168-8278(96)80280-7;
RX Aeshina Y., Enomoto N., Kuroaki M., Sakuma I., Izumi N., Marumo F.,
RA Sato C.;
RA "Sequential changes in full-length genomes of hepatitis B virus
RT accompanying acute exacerbation of chronic hepatitis B.";
RL J. Hepatol. 25:787-794(1996).
DR EMBL; D50518; BAA23442.1; -; Genomic DNA.
DR EMBL; D50517; BAA23435.1; -; Genomic DNA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNaPol_viral_C.
DR InterPro; IPR00201; DNaPol_viral_N.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00336; DNaPol_viral_C; 1.
DR Pfam; PF00242; DNaPol_viral_N; 1.

DR Pfam; PF00078; RVT_1; 2.
DR ProDom; PD000814; DNaPol_viral_C; 1.
SQ SEQUENCE 843 AA; 94462 MW; 1BCCF79639BB140D CRC64;
Alignment Scores: 2.79e-200 Length: 843
Pred. No.: 2802.00 Matches: 525
Score: 98.52% Conservative: 8
Percent Similarity: 97.04% Mismatches: 7
Best Local Similarity: 47.44% Indels: 2
Query Match: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x O42041_HPBVO (1-843)
Qy 1 CTCCACAACATTCCACCAGCTCTGCTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
Qy 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCCGACTTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGACTGGGGACCCCTGCACCGAACATGAGAACACCAACATCAGATTCTCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGACCCCTGCTGCTGTTTACAGCGGGGTTTTCTCGTTGACAAAGAACTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGACTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGGAGACCCACGCTGTTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisIleVal -Se 403
Qy 300 CTGCGCAAAATTCGAGTCCCAACCTCCAATCACTCACCAACCTCTGTCTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTCGCTATCTCGATGTGTCTCGGGGTTTTATCATATTTCTTCTTCTTCTTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCTCATCTTCTTGTGTTCTCTGACTTACCAAGTATGTTGCCGTTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCAGGAACATCAACCAACGACGCGGGGCATGCAAGACTGACACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSe 483
Qy 540 AAGGAAACTTACGTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCATCCCATCATCTCGGCTTCGACAGATTCCTATGGAGTGGGCTCTCTCTCTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTCAGTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACCTCTTATACCAATTTTCTTGTCTTGTGTTTGTGTTTGTGTTTGTGTT 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAACAACAAAGTTGGGGCTACTCCCTTAACCTCATCGGATATGTAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603

QY 900 GGGTACTTTACCGCAGGAACATATGTACTATAAACTCAAGCAATGTTTTCGAAAACTGCC 959
 Db |||||
 603 pGlyThrLeuProGlnGluHisIleValleuLysLeuGlnCysPheArgLysLeuPr 623
 QY 960 TCTAATAGACCTATTGATGTGAAGATGTCAAGAAATGTGGGTCTTTTGGCTTTGTC 1019
 Db |||||
 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
 QY 1020 TCCCTCTTTACACATGTGGCTATCTCCCTGATGCTTTATATGATGATGATATCAATC 1079
 Db |||||
 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
 QY 1080 TAAGCAGGCTTTCACTTTCTCCCAACTTACAAAGCCTTCTGTGTAAACAATATCTGAA 1139
 Db |||||
 663 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysGlnGlnTyLeuHi 683
 QY 1140 CCTTTACCCCGTTGCCGCAACGGTCCGTCTCTGCCAAGTGTGTGTGACGCAACCCC 1199
 Db |||||
 683 sLeuTyProValAlaArgGlnArgSerGlyValCysGlnValPheAlaAspAlaThrPr 703
 QY 1200 CACTGGATGGCGCTGGCCATAGGCCATCAGCGCATGGCTGGAACTTTCTGCTCCTCT 1259
 Db |||||
 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
 QY 1260 GCCGATCCCATCTGCGGAACCTCTAGCAGCTTGTGTTGTCGCGCGGTCTGGAGCAAA 1319
 Db |||||
 723 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
 QY 1320 ACTTATCGGAACCGAACCTCTGTGTTCTCTCTCGGAAATACACCTCTTTCCATGGCT 1379
 Db |||||
 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrple 763
 QY 1380 GTAGGGTGTGTCGCAACTGATCTGCGCGGAGCTCTTTGCTAGCTCCGCTCGGC 1439
 Db |||||
 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
 QY 1440 GCTGAATCCCGGAGCGACCCGTCTCGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
 Db |||||
 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuHi 803
 QY 1500 TCTGCGGTTCCGCGCACACCGGGCGACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
 Db |||||
 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
 QY 1560 TTCTCATCTGCGGACCGGTGTCACCTTCGCTTCACCTCTGCACGTCGATGGAGACCACC 1619
 Db |||||
 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 QY 1620 G 1620
 Db 843 O 843

Search completed: December 1, 2005, 23:59:04
 Job time : 772.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:15:45 ; Search time 106 Seconds

(without alignments)
5836.545 Million cell updates/sec

Title: US-10-761-006A-1

Perfect score: 5907

Sequence: 1 CTCACACATTCACACAG.....CCTCAGGCGCAGCGTGGAA 3215

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/csm2_1/USPTO.spool.p/US10761006/runat_01122005_113953_26882/app.query.fasta_1.3399
-DB=PIR -OPT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-OUTALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CGN 1 1 163 @runat_01122005_113953_26882 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2807	47.5	843	1 JDVLVR	DNA-directed DNA p
2	2789	47.2	843	2 S35527	DNA-directed DNA p
3	2750	46.6	843	1 S43491	DNA-directed DNA p
4	2734.5	46.3	842	1 JDVLVS	DNA-directed DNA p
5	2688	45.5	825	2 T13473	DNA-directed DNA p
6	2678.5	45.3	827	2 T13468	DNA-directed DNA p
7	2654	44.9	832	2 S71785	DNA-directed DNA p
8	2652	44.9	832	1 JDVLVB	DNA-directed DNA p
9	2651	44.9	832	1 S47406	DNA-directed DNA p
10	2643	44.7	832	1 JDVLVA	DNA-directed DNA p
11	2643	44.7	843	1 JDVLJ3	DNA-directed DNA p
12	2642	44.7	832	1 JDVLAI	DNA-directed DNA p
13	2637	44.6	832	1 S20757	DNA-directed DNA p
14	2628	44.5	832	2 S67505	DNA-directed DNA p

15	2628	44.5	843	1 JDVLJ1	DNA-directed DNA p
16	2623	44.4	832	1 JDVLCP	DNA-directed DNA p
17	2610	44.2	843	1 JDVLJ2	DNA-directed DNA p
18	2608	44.2	845	1 JDVLKS	DNA-directed DNA p
19	2607	44.1	832	1 S20752	DNA-directed DNA p
20	2607	44.1	845	1 JDVLVD	DNA-directed DNA p
21	2582	43.7	843	1 JQ2229	DNA-directed DNA p
22	2582	43.7	843	1 JDVLVH	DNA-directed DNA p
23	2230	37.8	750	1 S04570	gene X/C fusion pr
24	1719	29.1	884	1 JDVL7	DNA-directed DNA p
25	1716	29.1	884	1 JDVL59	DNA-directed DNA p
26	1714	29.0	556	1 JDVL64	DNA-directed DNA p
27	1708	28.9	884	1 JDVLW8	DNA-directed DNA p
28	1698	28.7	879	1 JDVLC	DNA-directed DNA p
29	1668	28.2	881	1 JDVLS	DNA-directed DNA p
30	1649	27.9	883	1 JDVLC2	DNA-directed DNA p
31	1472	24.9	400	1 SAVLA	large surface anti
32	1452	24.6	400	2 S35528	surface antigen -
33	1391.5	23.6	445	2 S43492	surface antigen -
34	1380	23.4	389	1 SAVLCP	large surface anti
35	1376	23.3	389	1 SAVLBH	large surface anti
36	1372	23.2	389	2 S67506	large surface anti
37	1369	23.2	389	2 S20745	surface antigen -
38	1368	23.2	400	1 JQ1575	major surface anti
39	1367	23.1	389	2 S32202	large surface anti
40	1363	23.1	389	2 S47407	surface antigen -
41	1363	23.1	389	2 S20753	surface antigen -
42	1361	23.0	400	1 SAVLVD	large surface anti
43	1356	23.0	389	2 S20749	surface antigen -
44	1355	22.9	384	2 T13469	large surface anti
45	1354	22.9	389	1 SAVLAH	large surface anti

ALIGNMENTS

RESULT 1

JDVLVR
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A00704
R:Ono, Y.; Onda, H.; Sasaki, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A:Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype A;Reference number: A93460; MUID:83168919; PMID:6300776
A:Accession: A00704
A:Molecule type: DNA
A:Residues: 1-843 <ONO>
A:Cross-references: UNIPROT:P03157; UNIPARC:UPI00001297FE
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1.13e-199 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLVR (1-843)

QY	4	CACAACTTCACCAAGCTCTGCTAGATCCAGGGGTGAGGGGCTATATTTTCCTGCTGG	63
Db	305	HisAenIIProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp	324
QY	64	TGGCTCAGTTCGGGAACAGTAACCCCTGCTCCGACTACTGCTCTCCATATCTCAAT	123
Db	325	TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAen	344
QY	124	CTTCTCGAGGAGCTGGGGAGCCCTGCAACCAATGAGAACACAACTCAGGATTCCTAGG	183
Db	345	LeuLeuGluAspTrpGlyProCysThrGluHisGlyHisAsnIleArgIleProArg	364

```
QY 184 ACCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACCAAGAACTCCTCAATACCGCA 243
Db |||
QY 365 ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGACCCACGTTGCTTG 302
Db |||
QY 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAATTCGCAGTCCCAACCTCCAACTCACTCAACCACTCTTGTCCTCCAAATTGTG 362
Db |||
QY 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCGCTGAGATGTCGTGGCGGTTTTATCATATTCTCTTCATCTCGTCTGCTAT 422
Db |||
QY 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
QY 423 GCCTCATCTCTTGTTGGTCTCTGGACTACCMAGGTATGTTGCCGGTTTGCTCTCTAC 482
Db |||
QY 444 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
QY 483 TTCAGGAACATCAACACAGACGCGGGCCATGCAAGACCTGCACGACTCCTGCTCAAG 542
Db |||
QY 464 rSerArgAsnLleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTAGCTTTCCCTCTTGTTGCTGTACAAACCTTCGGACGGAAACTGCACCTTGTA 602
Db |||
QY 484 gAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 504
QY 603 TTCCCATCCCATCATCTCGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCTAGTCGTT 662
Db |||
QY 504 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
QY 663 TCTCTCGCTCAGTTTACTAGTGCATTGCTCAGTGGTTCGTAGGGCTTTCCTCCCACTG 722
Db |||
QY 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544
QY 723 TTTGGCTTTCAGTTATATGATGTGTGTTATTTGGGGGGGAAAGTCTGTACAACTTTGA 782
Db |||
QY 544 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTGTCCTTTGGGTATACATTTAAACCTTAA 842
Db |||
QY 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAACCAAAAGTGTGGGGTACTCCCTTAACTTCATGGGATATGTAAATTCGAAAGTTGGG 902
Db |||
QY 584 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 604
QY 903 TACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACTGCCTGT 962
Db |||
QY 604 yThrLeuProGlnGlnHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AAATAGACCTATTGATTGGAAAGTATGTCAAAGAAATTGTGGGTCTTTTGGGCTTTGCTGC 1022
Db |||
QY 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
QY 1023 CCCTTTTACCAATATGCGCTATCTGCTTGATGCGCTTTTATATGTCATGTATACAACTTAA 1082
Db |||
QY 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 664
QY 1083 GCAGGCTTTCACCTTTCTCGCCAACTTACAAAGCCCTTTCTGTGTAAACAATATCTGAACCT 1142
Db |||
QY 664 gGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
QY 1143 TTACCCGTTGCCCGGCAAGGTCGGTCTCTGCGCAAGTGTGTGCTGACGCAACCCCCAC 1202
Db |||
QY 684 uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
QY 1203 TGGATGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTGCTGCC 1262
Db |||
QY 704 rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
```

```
QY 1263 GATCCATACTCGGGAACCTCTAGCAGCTTGTTTGTCTCCAGCCGCTCTGGAGCAAAACT 1322
Db |||
QY 724 oIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
QY 1323 TATCGGAACCGCAACTCTGTTGTCTCTCTCGGAAATACACCTCTTTCCATGGCTGCT 1382
Db |||
QY 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
QY 1383 AGGGTGTGCTGCGCAACTGGATCTGCGCGGACGCTCTTTGTCTACGTCCTCCGTCGCGCT 1442
Db |||
QY 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLe 784
QY 1443 GAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCACT 1502
Db |||
QY 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLe 804
QY 1503 GCGGTTCGCGCGACACCGGGCGGACCTCTTTACGGGCTCTCCCGTATGTGCTTC 1562
Db |||
QY 804 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 824
QY 1563 TCATCTGCGCGACCGGTGTCACCTTCGCTTCACCTCTGCACGTCGCGATGAGACACCG 1620
Db |||
QY 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
RESULT 2
S35527
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype adr)
C;Species: hepatitis B virus, HBV
A;Variety: subtype adr
C;Date: 09-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: S35527
R;Mukaide, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiji, K.
Nucleic Acids Res. 20, 6105, 1992
A;Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and
A;Reference number: S35527; MUID:93096607; PMID:1461746
A;Accession: S35527
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-843 <WUK>
A;Cross-references: UNIPROT:Q81107; UNIPARC:UPI00000BED31; EMBL:D12980; NID:g221500; PIDN
C;Experimental source: subtype adr
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C;Genetics:
A;Introns: 345/3
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 2,44e-198 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x S35527 (1-843)
QY 1 CTCACAACATTCACCAAGCTCTGCTAGATCCAGGGTGAGGGCGCTATATTTCTCTGC 60
Db |||
QY 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTTCGCGAACAGTAAACCTGTTCGAGCTACTGCTCTCCCATATCGTC 120
Db |||
QY 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTGAGACTGGGGACCTTCGACCGCAACATFGAGAAACACAACTACAGATTCTCT 180
Db |||
QY 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTTACAGCGGGGTTTTCTCGTTGACAAAGATCTCTCAATATACC 240
Db |||
QY 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
```

```
QY 241 GCAGAGTCTAGACTCTG-CTGGACTTCTCTCAATTTCTAGGGGAGCACCCAGTGTTC 299
Db      |||||
QY 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCAAAATTCGCGAGTCCCAACCTCCAACTCACTCAACCAACCTTCTGTCTCTCAATTT 359
Db      |||||
QY 403 rTrpProlysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 CTCTCGGTATCGCTGGAGTGTCTGCGCGGTTTATCATATTCCTTCTTCATCTCGTGC 419
Db      |||||
QY 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeProLeuHisProAlaAl 443
QY 420 TATGCTCATCTTCTTGTGTTCTCTCGACTACCAAGATTATGTCCTGCTTCCTC 479
Db      |||||
QY 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyValAlaAla-gLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGACGCGGCCCATGCAAGACCTTCGACGGAACCTGCCTC 539
Db      |||||
QY 463 rThrSerArgAsnIleAsnTyHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGAAACTCTACGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db      |||||
QY 483 rArgAsnLeuTyValSerLeuLeuValTyHisThrPheGlyArgLeuLeuHisLe 503
QY 600 GTAATCCCATCCATCATCTCTGGCTTCGCAAGATTCTATGGAGTGGGCTCAGTCC 659
Db      |||||
QY 503 uTySerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGGCTCAGTTTACTAGTGCATTTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 719
Db      |||||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
Db      |||||
QY 543 sCysLeuAlaPheSerTyMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACTCTTATACCAATTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
Db      |||||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGCTTACTCTTAACTTCATGGCATATGTAATTTGGAAGTTG 899
Db      |||||
QY 583 oAsnLysThrLysA-grtpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 603
QY 900 GGGTACTTTACGCGAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db      |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATGGAAAGTATGTCAAGAATTGGGTCCTTTTGGGCTTTGC 1019
Db      |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCTCTTTTACCAATGTGGCTATCTCTGCTGTATGCTTATATGATGTATATCAATC 1079
Db      |||||
QY 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db      |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyHisAlaPheLeuCysGlnGlnTyHisLeuHi 683
QY 1140 CTTTACCGGTTGCGCGCAACGGTCCGGTCTCTCCCAAGTTTGTCTGACGCAACCC 1199
Db      |||||
QY 683 sLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
QY 1200 CACTGTAGTGGGCTTCGCCATAGGCCATGCGCTGCTGGAACCTTTCTGGCTCTCT 1259
Db      |||||
QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATCTACGCGAACTCTAGCAGCTTCTTTTGTCTCGCAGCGGCTCTGGAGCAA 1319
Db      |||||
QY 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTy 743
QY 1320 ACTTATCGGAACCGAACACTCTGTGTCTCTCTCGGAATATACCTCTCTTCATGGCT 1379
```

```
Db      |||||
QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAACATGATCTCTGCGGGGAGAGTCTCTTGTCTACGTCCCGTCCGC 1439
Db      |||||
QY 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 783
QY 1440 GCTCAATCCCGCGAGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCTTCA 1499
Db      |||||
QY 783 aLeuAsnProAlaAspAspProSerArgLysGlyLeuTyArgProLeuLeuHi 803
QY 1500 TCTCGCTTCGCGCGACGACGCGCGCTCTCTTTTACGCGGTCTCCCGTATGTGCC 1559
Db      |||||
QY 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGGTGTGCATCTCGCTTACCTCTGACGCTCGCATGGAGACCACC 1619
Db      |||||
QY 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
Db      |||||
QY 843 o 843

RESULT 3
S43491
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr)
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 07-Sep-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C:Accession: S43491
R:Loncarevic, I.F.; Zentgraf, H.; Schroeder, C.H.
Nucleic Acids Res. 18, 4940, 1990
A:Title: Sequence of a replication competent hepatitis B virus genome with a preX open
A:Reference number: S12598; MUID:90370503; PMID:2395664
A:Accession: S43491
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-843 <LON>
A:Cross-references: UNIPROT:Q67952; UNIPARC:UPI00000EE7B3; EMBL:X52939; NID:g457780; PFI
A:Experimental source: subtype adr
C:Genetics:
A:Gene: P; pol
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 1.89e-195 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 1

US-10-761-006A-1 (1-3215) x S43491 (1-843)
QY 1 CTCGACACATTTCCACCAAGCTCTCTAGATCCAGGTCGAGGGCTATATTTCTCTCC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerLysSerGlyProLeuPheProCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTCTGCTCCGACTACTCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysThrGluHisGlyGluHisAsnIleArgIleVal 343
QY 121 AATCTTCTCAGGACTGGGGACCTTCGACCGAACATGGAGAACACAAATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGAGCCCTCTCGTGTTCACGCGGGGTTTCTCGTTGACAGAAATCCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-CTGGACTTCTCTCAATTTCTAGGGGAGCACCCAGTGTTC 299
```

Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
Qy 300 CTGGCCAAAATTCGAGTCCCAACCTCCAATCACTCAACCAACCTCTTGTCTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGGCTATCGCTGGATGTCTCGCGCGTCTTTATCATATTCCTCTTCATCTCGTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGTCTCTCGGACTACCAAGGTATGTGCGGTCTTGTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACACCGACGCGGCCCATGCAAGACCTTCGACGACTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
Qy 540 AAGGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGAGCGAAATGCACTT 599
Db 483 rLysHisLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
Qy 600 GTATTCCCATCCCATCATCTCGGCTTCGCAAGATTCTATGGAGTGGSCCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTGTTCAGTGGTTCGTAGGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTCAGTTATATGGATGATGTGGTATTTGGGGCGAAGCTGTACAAATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
Qy 780 TGAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheMetCysLeuGlyIleHisLeuAsnPr 583
Qy 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
Qy 900 GGGTACTTTACCGAGGACATATGTACTAAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGCAAGAATTCTGGGTCTTTTGGGCTTGC 1019
Db 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACAAATGTGGCTATTCCTGCTTTCATGCTTTATATGATGATGATACAAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAACGAGCTTTCACCTTCGCGCAACTTACAGCGCTTCTGTGTAAACAATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuGlyGlnTyrLeuAs 683
Qy 1140 CTTTTACCCCTGTCGCGGCAACGCTCGGTCTCTGCCAAGTGTTCGTACGCAACCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGGATGGGGCTTGGCCATAGCCCATCAGCGCATGGCTGGAAACCTTTCGGCTCCTCT 1259
Db 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCATCCATCTCGGAACTCCTAGCAGCTTGTGTCGCGCGCGGTCTGAGCAAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaTh 743
Qy 1320 ACTTATCGGAACGACAACTCTGTGTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763

Qy 1380 GCTAGGTGTCTCTGCCAACTGGATCTCTCGCGGAGCGTCTTTGTCTACGTCCCGTCGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCGCGGACGACCGCTCTCGGGCGCGTTCGGGGCTCTACCGTCCCTTCTTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
Qy 1500 TCTGCCGTTCGCGCGCACCGACGCGGCGACCTCTCTTTAGCGGTCTCCCGTATGTGCC 1559
Db 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTCGCGACCGCTGTGCATCTTCGCTTCACCTCTGCACCTCGCATGGACACACC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
Qy 1620 G 1620
Db 843 o 843
RESULT 4
JDVLVS
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adr, mutant)
C:Species: hepatitis B virus, HBV
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S04568
R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.
Nucleic Acids Res. 17, 2124, 1989
A:Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp
A:Reference number: S04568; MUID:89183619; PMID:2928116
A:Accession: S04568
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-842 <RHO>
A:Cross-references: UNIPROT:P31870; UNIPARC:UPI00001297FB; EMBL:X14193
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 2,65e-194 Length: 842
Score: 2734.50 Matches: 517
Percent Similarity: 96.86% Conservative: 7
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.29% Indels: 3
DB: 1 Gaps: 1
US-10-761-006A-1 (1-3215) x JDVLVS (1-842)
Qy 1 CTCACAACATTCACCAAGCTCTGTAGATCCCGGGTGAGGGGCTATATTTCTCTGC 60
Db 304 LeuHisIleSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
Qy 61 TGTGGTCCAGTTCGGAACAGTAAACCTGTGTCCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTCTTCGAGGACTGGGACCCCTGCACCAGACATCGAGACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyLeuHisAsnIleArgGilePro 363
Qy 181 AGGACCCCTCTCTGTGTACAGCGGGGTTTTCTCGTTGACAAGAAATCTCTCACAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG- GTGGACTTCTCTCAATTTTCTAGGGGGGACACCCAGTGTCT 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGCCCAAAATTCGAGTCCCAACCTCCAATCACTCAACCAACCTCTTGTCTCTCCAAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423

QY 360 GTCCTGGCTACGCTGGATGTCCTGGCGGTTTATCATATTCCTCTTCATCTGCTGC 419
Db |||||
QY 423 userftrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATCCCTCATCTCTCTGTTGGTCTCTTGGACCTACCAAGGTATGTTGCCGGTTTGTCTCT 479
Db |||||
QY 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaargLeuSerSe 463
QY 480 TACTTCAGGAACTACCAACACACAGCAGCGGGCCATGCAAGACCTGCACGACTCTGCTCT 539
Db |||||
QY 463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisaspSerCysSe 483
QY 540 AAGGAAACTCTACCTTTCCCTCTGTTGCTGTACAAACCTTCGACGGAACTGCACCT 599
Db |||||
QY 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCTCATCTCCATCTCTGGGCTTTGCAAGATTCCTATGGAGTGGGCTTCAGTCC 659
Db |||||
QY 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr 523
QY 660 GTTTCCTCTGCTCAGTTTACTAGTCATTTGTCAGTGTGTTGCTAGGGCTTTCCCCCA 719
Db |||||
QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTTGTTGGGGCGAGTCTGTACACATCT 779
Db |||||
QY 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCTCTTTTACCTTATTACCAATTTCTTTGTTGTTGGTATACATTTAAACCC 839
Db |||||
QY 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAAGCTTGGGCTACTCCCTTAACTTCATGGATATGTAATTTGGAAGTTG 899
Db |||||
QY 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGAGGACATATGTACTTAAACTCAGCAATGTTTTCGAAACCTGCC 959
Db |||||
QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAATAAGACTTATGATGGAAGATGTCAAGAAATCTGGTGTCTTTGGCTTTGC 1019
Db |||||
QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACAAATGTGCTACTCTGCTGCTGCTTATGATGCTTATATGATGATATACATC 1079
Db |||||
QY 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db |||||
QY 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTACCCCTTCCCGGCAACGGTCCGGTCTCTGCAAGTGTTCGTGACGCAACCCC 1199
Db |||||
QY 683 sLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGTGATGGGCTTGGGCATAGGCATCAGCGATGCTGGAACCTTTCTGGCTCTCT 1259
Db |||||
QY 702 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCCATATCTCGGAACTCTTAGCAGCTTGTGTTGCTCGCAGCCGGTCTGGAGCAA 1319
Db |||||
QY 722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
QY 1320 ACTTATCGGAACCGACACTCTGTGCTCTCTCGGAAATACACCTCTTTTCCATGGCT 1379
Db |||||
QY 742 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGGTGTGTCGCAACTGATCTCGCGGGGACGTCCTTTGTGTACGTCCTCGGCTGGC 1439
Db |||||
QY 762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGGGCGCTTTGGGGCTCTACCGTCCCTCTTCTCA 1499

Db 782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuLeuArgProLeuLeuHi 802
QY 1500 TCTGCGCTTCCGCGCCACACACGCGGCGCCTCTCTTTAGCGGTCTCCCGGTATGTC 1559
Db |||||
QY 802 sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTTCTCATCTGCGGACCGGTGTCACCTTCGCTTTCACCTCTGACGTGCGATGGAGACCC 1619
Db |||||
QY 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
QY 1620 G 1620
Db |||||
Db 842 o 842
RESULT 5
T13473
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 29Y11HCC)
C;Species: hepatitis B virus, HBV
A;Variety: isolate 29Y11HCC
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T13473
R;Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A;Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
A;Reference number: Z17684; MUID:99129050; PMID:9930189
A;Accession: T13473
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-825 <TAK>
A;Cross-references: UNIPROT:Q9Y2S3; UNIPARC:UPI00000F4BF5; EMBL:AB014388; NID:G3582381;
A;Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HCC
C;Genetics:
A;Gene: P
A;Introns: 303/3
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase
Alignment Scores:
Pred. No.: 7.33e-191 Length: 825
Score: 2688.00 Matches: 504
Percent Similarity: 95.92% Conservative: 13
Best Local Similarity: 93.51% Mismatches: 21
Query Match: 45.51% Indels: 2
DB: Gaps: 0
US-10-761-006A-1 (1-3215) x T13473 (1-825)
QY 7 AACATTCACCAAGCTCTGCTAGATCCAGGTGAGGGCTATATTTCTGCTGCTGG 66
Db |||||
QY 288 HisLeuSerThrThrLysArgGlnSerSerSerGlyHisAlaValGluThrCysTrpTrp 307
QY 67 CTCAGTTCGGACAGTAACCTGTTCCGACTACTGCTCTCCCATATGCTCAATCTT 126
Db |||||
QY 308 LeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeu 327
QY 127 CTCGAGACTGGGACCTCGACCGAATATGGAGAACCAACATCAGATTCCTTAGGACC 186
Db |||||
QY 328 LeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIleProArgThr 347
QY 187 CTTGCTCGTGTATAGCGGGGTTTTCTCGTTGACAAGAATCTCTCAATACCGCAGAG 246
Db |||||
QY 348 ProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGlu 367
QY 247 TCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGGACCCACGCTTCTGCGCC 305
Db |||||
QY 368 Ser***LeuValValAspPheSerGlnPheSerArgGlySerThrGlnVal-SerTrpPr 387
QY 306 AAAATTCCGAGTCCCAACCTCCAACTCACCTCACCAAGCTCTTGTCTCCCAATTTGCTGT 365
Db |||||
QY 387 oLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTr 407
QY 366 GCTATCGCTGGATGTGTCTGGCGGTTTTTATCATATTTCTCTTCTCATCTGCTGATGCC 425


```

Db      407 pLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetPr 427
      |||
QY      426 TCATCTTCTTGTGTCTCTTCGGAACCAAGGTATGTTGGCCGTGTCTCTACTTTC 485
      |||
Db      427 ofHisLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSe 447
      |||
QY      486 CAGGAACATCAACCAACGACGCGGCCATGCAAGACCTGCAAGCTCTCTGCTCAAGGAA 545
      |||
Db      447 rArgAsnIleAen**GinHisGlyThrMetGlnAspLeuHisAspSerCysSerLysHi 467
      |||
QY      546 ACTTACGTTTCCCTCTCTGCTGTACAAACCTTCGACGGAACCTGCACTGTATTC 605
      |||
Db      467 sLeuTyrValSerLeuLeuLeuTyrIlePheGlyArgLysLeuHisLeuTyrSe 487
      |||
QY      606 CCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTTCT 665
      |||
Db      487 rHisProIle**LeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLe 507
      |||
QY      666 CTGCTCTCAGTTTACTAGTGCATTTGTTTCACTGGTTCGTAGGCTTTCCTCCCACTGTTT 725
      |||
Db      507 uLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLe 527
      |||
QY      726 GCCTTTCAGTTATGATGATGTGGTATTTGGGGGGAAGTCTGTACAACTCTTGATC 785
      |||
Db      527 uAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeuGluAl 547
      |||
QY      786 CTTTCTTACCTTATACCAATTTCTTTTGTCTTGGGTATACATTTAAACCTCAATAA 845
      |||
Db      547 aLeuTyrThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLy 567
      |||
QY      846 AACCAAACTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTTGGAAGTTGGGGTAC 905
      |||
Db      567 sThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyrGlyTh 587
      |||
QY      906 TTTACGCGAGGAACATATGTAATAACTAAGCAATGTTTTCGAAACCTCGCTGTAAA 965
      |||
Db      587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
      |||
QY      966 TAGACCTATTGATGGAAGTATGCAAGAATTGGGTCTTTTGGGCTTTTGGCTGCTGCC 1025
      |||
Db      607 nArgProIleAspTyrLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr 627
      |||
QY      1026 TTTTACCAATGTGGCTATCTCGCTTGATGCTTTTATATGATGATGATATCAATCTAAGCA 1085
      |||
Db      627 oPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLysGl 647
      |||
QY      1086 GGTCTTTCATCTTCTCGCCAACTTACAGGCTTTCGTGTAAACAATATCTGAACCTTTA 1145
      |||
Db      647 nAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLeuTy 667
      |||
QY      1146 CCGCGTTGCCGCAACGTCGCTCTCGCAAGTGTTCGTGACGCAACCCCACTGCG 1205
      |||
Db      667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
      |||
QY      1206 ATGGGCTTGGCCATAGGCGCATGACGCGATGGCTGGAACCTTTCTGGCTCCTCTGCGCAT 1265
      |||
Db      687 yTyrGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
      |||
QY      1266 CCATATCGGGAACCTCTAGCAGCTTGTGTTGCTCGACCGGCTGAGAGCAAACTTAT 1325
      |||
Db      707 eHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
      |||
QY      1326 CGGAACCGCACTCTGTGTCTCTCTCGGAATACACTCTTCTCCATGCTGCTAGG 1385
      |||
Db      727 eGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTyrProLeuLeuGl 747
      |||
QY      1386 GTGTGCTGCGCAACTGATCTCGCGGGGAGCTCTTTGTCTAGTCCCGCTCGCGCTGAA 1445
      |||
Db      747 yCysAlaAlaAenTyrIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAs 767
      |||
QY      1446 TCCCGGGAAGACCCGCTCGCGGGCGGTTTGGGGCTCTACCGTCCCTCTTCATCTGCC 1505
      |||

```

```

Db      767 nProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuPr 787
      |||
QY      1506 GTTCGGCGCCGACCAACGGGCGCACCTCTCTTTACGGGTCTCCCGTATGCTCTTCTCA 1565
      |||
Db      787 oPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHi 807
      |||
QY      1566 TCTCGCGGACGCTGTGCACTTCGCTTACCTCTGCACTGCGACGTCGATGAGACACCG 1620
      |||
Db      807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTyrArgProPro 825
      |||
RESULT 6
Tl3468
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (isolate 09D09HCC)
C:Species: hepatitis B virus, HBV
A:Variety: isolate 09D09HCC
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: Tl3468
R:Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.
Arch. Virol. 143, 2313-2326, 1998
A:Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin
A:Reference number: Z17684; MUID:99129050; PMID:9930189
A:Accession: Tl3468
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-827 <TAK>
A:Cross-references: UNIPROT:Q9YZU3; UNIPARC:UPI00000F479C; EMBL:AB014368; NID:g3551304;
A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC
C:Genetics:
A:Gene: P
A:Introns: 303/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.:      3,7e-190      Length:      827
Score:          2678.50      Matches:    504
Percent Similarity: 95.91%      Conservative: 12
Best Local Similarity: 93.66%      Mismatches:  18
Query Match:    45.34%      Indels:      5
DB:             2           Gaps:        1

US-10-761-006A-1 (1-3215) x Tl3468 (1-827)

QY      19 AGCTCTGCTAGATCCACGGGTGAGGG-----CCTATATTTTCTGCTGGTGCTC 69
      |||
Db      291 ThrSerLysArgGlnSerSerGlyHisAlaValGluLeuHisProCysTyrTrpLeu 310
      |||
QY      70 GAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
      |||
Db      311 GlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsnLeuLeu 330
      |||
QY      130 GAGGACTGGGGACCTCGCACCGAACATGGAGAACACACATCAGGATTCCTAGGACCCCT 189
      |||
Db      331 GluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgThrPro 350
      |||
QY      190 GCTCGTGTGTACAGGGGGGTTTTCTCGTTTGAAGAATCTCTCACAAATCCCGCAGAGTCT 249
      |||
Db      351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370
      |||
QY      250 AGACTCTG-GTGAGCTTCTCTCAATTTTCTAGGGGAGCACCCACGCTGTTCTTGGCCAAA 308
      |||
Db      371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr***Val-SerTyrProLy 390
      |||
QY      309 ATTGCGAGTCCCAACCTCCAATCTACACCACTCTTGTCTCTCCAAATTTGTCCTGGCT 368
      |||
Db      390 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTyrPle 410
      |||
QY      369 ATCGCTGGAATGTCTGCGCGCGTTTTTATCATATTCCTTCTTCATCTCTGCTGCTATGCTCA 428
      |||
Db      410 uSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMetProHi 430
      |||
QY      429 TCTTCTTGTGTGTTCTTCTGGAATACCAAGGTATGTTGGCGTGTGCTCTCTACTTCTCCAG 488
      |||

```



```
Db 430 sLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSerAr 450
Qy 489 GAACATCAACACACAGCAGCGGCCATGCAAGACCTGCACGACTCTGCTCAAGGAACCT 548
Db 450 GAsn**AsnTyrGlnHisGlyThrMetGlnAspLeuHisGluSerCysSerArgAenLe 470
Qy 549 CTAGCTTCTCCCTCTGTTGCTGTACAAACCTTCGGACGGAACCTGACCTGATATCCCA 608
Db 470 uTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLeuHisLeuTyrSerHi 490
Qy 609 TCCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCGTTTCCTCT 668
Db 490 sProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLeuLe 510
Qy 669 GGCTCAGTTACTAGTGCATTTCTTCAGTGGTTCGTAGGCTTTCCTCCCACTGTTGGC 728
Db 510 uAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHisCysLeuAl 530
Qy 729 TTTTCAGTTATATGATGATGTTGTTGGGGCGAAGTCTGTACAAATCTTCAGTCCCT 788
Db 530 aPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
Qy 789 TTTTACCTCTATTACCAATTTCTTTGTCTTTGGGTATACATTTAAACCTTAATAAAC 848
Db 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyLysLeuAsnProHisLysTh 570
Qy 849 CAACGTTGGGGCTACTCCCTTAATCTTATGGGATATGTAATCGAAGTTGGGGTACTTT 908
Db 570 xLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTyrGlyThrLe 590
Qy 909 ACCGACGACATATTGTTACTTAAACCTCAAGCAATGTTTCGAAAACTGCTGTAAATAG 968
Db 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
Qy 969 ACCTATTGATTGGAAGATGTTCAAGAAATGTCGGTCTTTTGGGCTTTGTCGCCCTTT 1028
Db 610 sProIleAspTyrLysValCysGln**IleValGlyLeuLeuGlyPheAlaAlaProPh 630
Qy 1029 TACAATATGGGCTATCTGCTTGTATGCTTCTTATATGATGATATCAATCTTAAGCAGGC 1088
Db 630 eThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAlaLysGlnAl 650
Qy 1089 TTTTCACTTCTCGGCAACTTACAAGGCTTCTGTGTAAACATATCTGAACCTTTACCC 1148
Db 650 aPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHisLeuTyrPr 670
Qy 1149 CGTTGCCCGCAACGGTCCGCTCTGCAAGTGTGTCGACGCAACCCCACTGGATG 1208
Db 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Qy 1209 GGGCTTGGCCATAGGCCATCAGCGATGGCTGGAACCTTTCTGCTCTCTGCGGATCCA 1268
Db 690 pGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Qy 1269 TACTCGCGAACTCTAGCAGCTTGTGTTGCTCGGACGCGTCTCGAGCAAACTATCGG 1328
Db 710 sThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Qy 1329 AACCGACAACCTGTGTGCTCTCTCGGAATACACCTCTCTTCATGCTGTAGGGTG 1388
Db 730 yThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProThrLeuLeuGlyCy 750
Qy 1389 TGCTGCCAACTCGGATCTCTGCGCGGACGTCTTTGTCTACGTCCTCGCGCTGAATCC 1448
Db 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAsnPr 770
Qy 1449 CGCGACGACCGCTCTCGGCGCGTGTGGGGCTCTACGTCCTCCCTCTTCTCATCTGCGGTT 1508
Db 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuProPh 790
Qy 1509 CGGCGCGACGCGGGCCACTCTCTTTAGCGGTCTCCCGGTATGCGCTCTCATCT 1568
Db 790 eArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHisLe 810
```

```
Qy 1569 GCCGGACCGTGTGCACCTTCGCTTACCTCTGCACGTCGCGATGAGAGACCACCG 1620
Db 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
```

RESULT 7

S71785
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, isolate pati
C;Species: hepatitis B virus, HBV
A;Variety: subtype ayw, isolate patient C1005
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S71785
R;Preisler-Adams, S.; Schlaver, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
submitted to the EMBL Data Library, March 1993
A;Description: Identification and sequence analysis of hepatitis B virus DNA in immunol
A;Reference number: S32202
A;Accession: S71785
A;Molecule type: DNA
A;Residues: 1-832 <PRE>
A;Cross-references: UNIPROT:Q8JN11; UNIPROT:Q9IF40; UNIPROT:Q96846; UNIPROT:Q67892; UNI
PROT:Q9QAG0; UNIPROT:Q9QAF3; UNIPROT:Q8QXQ1; UNIPROT:O11885; UNIPROT:O56654; UNIPROT:O5
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.:	2,42e-188	Length:	832
Score:	2654.00	Matches:	492
Percent Similarity:	95.56%	Conservative:	25
Best Local Similarity:	90.94%	Mismatches:	23
Query Match:	44.93%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006a-1 (1-3215) x S71785 (1-832)

```
Qy 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGGAGGCGCTATATTTTCCTGC 60
Db 293 LeuHisAsnPheProAsnSerAlaArgSerGlnGlyGluArgProAlaPheProCys 312
Qy 61 TGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCGAGTACTGCTCTCCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
Qy 121 AATCTTCTCAGAGCTGGGGACCTCGACCGAATCATGAGAACACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
Qy 181 AGGACCCCTGCTGTGTACAGCGGGGTTTTTCTCGTTGACAGAATCTCTACAATACC 240
Db 353 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 372
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCCACGTGTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se 392
Qy 300 CTGGCCAAATTCGAGTCCCAACCTCCAACTCATCTCACCACCTCTGCTCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
Qy 360 GTCTGGCTATCGCTGGATGTCGTGGCGGTTTTTATCATATTCCTCTCTCATCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaSe 432
Qy 420 TATGCTCTCATCTCTTCTGTTGCTTCTCTGAGCTACCAAGGTATGTTGCCGTTTGTCTC 479
Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
Qy 480 TACTTCCAGGAATCAACACCGACGACGGGGCCATGCAAGACTGACGACCTGCTCTC 539
Db 452 rAsnSerArgIlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
Qy 540 AAGGAAACTCTAGCTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACTGCACTT 599
```

Db 472 rAAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492

QY 600 GTATTCCCATCCCATCATCTCGGCTTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659

Db 492 utySerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512

QY 660 GTTTCCTCGGCTCAGTTACTAGTGCCTATTTGTTTCAGTGGTTCGTAGGCGTTCCTCCCA 719

Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532

QY 720 CTGTTTGGCTTTTCAGTTATATGATCATCTGGTATTGGGGCGCAAGTCTGTACCAATCT 779

Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552

QY 780 TGAGTCCCTTTTACCTTATACCAATTTCTTTGTTCTTTGGGTATACATTTAAACCC 839

Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuSerLeuGlyIleHisLeuAsnPr 572

QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTGG 899

Db 572 oAsnLysThrLysArgTyrGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy 592

QY 900 GGGTACTTTTACCGACGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACCTGCC 959

Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612

QY 960 TGTAAATAGACCTTATGATTGGAAGTATGTCAAAGAATTGTGGGTCTTTTGGGCTTTGC 1019

Db 612 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632

QY 1020 TGCCCTTTTACACATGTTGGTATCTCTGCTTGATGCTTTATATGATGATATACAATC 1079

Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652

QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139

Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672

QY 1140 CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCCAAGTGTGTTGTCGACCAACCC 1199

Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692

QY 1200 CACTGGATGGGCTTGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259

Db 692 oThrGlyTyrGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712

QY 1260 GCGATCCATATGCGGAACCTCTAGCAGCTTGTGTTGTTGTCGCGCGGTCTGGAGCAA 1319

Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732

QY 1320 ACTTATCGGAACCGACAACCTCTGTTGCTCTCTCGGAATACACCTCTTTCCATGGCT 1379

Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752

QY 1380 GCTAGGGTGTGTCGAACCTAGTCTCGCGGAGCTCTTTGCTAGCTCCGCTCGGCG 1439

Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772

QY 1440 GCTGAATCCCGCGAGACCGCTCTCGGGCGCGTTTGGGCTCTACCGTCCCTCTTCCA 1499

Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyIlePheArgProLeuLeuAr 792

QY 1500 TCTGCGGTTCCGCGCACCAACCGGCGCACCTCTCTTTACGGGCTCTCCCGTATGTGCC 1559

Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812

QY 1560 TTCTCATCTCCGGACCGGTGTGCACTTCGCTTACCTCTGCGAGCTGCGATGGAGACCAC 1619

Db 812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832

QY 1620 G 1620

Db 832 o 832

RESULT 8

JDVLVB

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, strain pH320
C:Species: hepatitis B virus, HBV
A>Note: host Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A00703
R:Richko, V.; Pushko, P.; Dreilina, D.; Pumpen, P.; Gren, E.
FEBS Lett. 185, 208-212, 1985
A>Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.
A:Reference number: A05237; MUID:85204397; PMID:3996597
A:Accession: A00703
A:Molecule type: DNA
A:Residues: 1-832 <BIC>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI0000171088; GB:X02496; NID:G62280; PIDN:C
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:

Pred. No.: 3,4e-188 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x JDVLVB (1-832)

QY 1 CTCCACACATTCACCAAGCTCTGTAGATCCAGGGGTAGGGGCTATATTTCCTGTC 60

Db 293 LeuHisAsnLeuProProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312

QY 61 TGGTGGCTCCAGTTCGCGAACAGTAAACCTGTTCCGACTACTGCTCTCCATATCTGC 120

Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332

QY 121 AATCTTCTCGAGGACTGGGGACCCCTGACCCGAAACATGGAGAACACACATCAGGATTCCT 180

Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352

QY 181 AGGACCCCTGCTGTTTACAGGGGGGTTTTCTCGTTGACAAAGATCCTCACAAATACC 240

Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372

QY 241 GCAGAGTCTAGACTCTG- GTGAGCTTCTCTCAATTTCTAGGGGAGCACCCACGCTGTC 299

Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgLysAsnTyrArgVal-Se 392

QY 300 CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACCACTCTTGTCTCCAAATTT 359

Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412

QY 360 GTCTGGCTATCGCTGGATGTGTCTGGCGGCTTTTATCATATTCCTCTTCATCTGCTGC 419

Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432

QY 420 TATGCTTCATCTCTTGTGTTCTCTCGAGTACCAGGATATGTTGCCGTTTGTCTCTC 479

Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452

QY 480 TACTTCCAGGAACATCAACCAACCAACCGGGCCATGCAAGACCTGACGACTCTCTGCTC 539

Db 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472

QY 540 AAGGAAACTTACGTTTCCCTCTGTTCTGTACAAACCTTCGGACCGGAACCTGCATT 599

Db 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492

QY 600 GTATTTCCCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659

Db 492 utySerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512

```
QY 660 GTTCTCTCTGGCTCAGTTTACTAGTCCATTTGTTCTAGTGGTTCGTAGGCTTTCCCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaIleSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATACCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 839
Db 552 uGluSerLeuPheThrAlaValThrAsnLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGATATGATTAATTCGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTySerLeuHisPheMetGlyTyValIleGlyCysTy 592
QY 900 GGGTACTTTACCGCAGGAACATATGTTACTAAACCTCAAGCAATGTTTTCGAAACCTGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATPAGACCTATTGATTTGGAAGATGTCAAAGAAATTTGTGGTCTTTTGGCTTTGC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TSCCCCTTTTACACNATGTGGCTATCTGCTTGTATGCTCTTTATATGATGATATACATC 1079
Db 632 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 652
QY 1080 TAAGCAGGCTTTCACCTTCTCGCAACTATACAGGCTTCTGTGTGTAAACATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLeuAs 672
QY 1140 CCTTTACCCGTTGCGGCAACGTCGCTCTCTGCAAGTGTGCTGACGCAACCC 1199
Db 672 nLeuTyProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGCCATAGCCATCAGCCATGCTGGAACCTTTCTGGCTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
QY 1260 GCGCATCATCTGCGCAACTCTTAGCAGCTTGTGTTGCTCGCAGCGGCTCGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGAACCGCAACTCTGTGCTCTCTCGGAATATACACCTCTTTCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyThrSerTyProTrpLe 752
QY 1380 GCTAGGTGTGCTGCCAATCGATCTCTGCGGGGACGTCCTTTGCTACGTCGCTGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
QY 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGCGGCTTTCGGGCTCTACCGTCCCTTCTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCCGTTCCGGCGGACGCGGCGCACCTCTCTTTACGCGGTCTCCCGGTATGTC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCTGCTTACCTCTGACGTCGATCGGACGACCC 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832
```

RESULT 9

S47406

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw4)

C;Species: hepatitis B virus, HBV

A;Variety: subtype ayw4

C;Date: 23-Nov-1994 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: S47406
R;Pluclenniczak, A.
submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isola
A;Reference number: S47404
A;Accession: S47406
A;Molecule type: DNA
A;Residues: 1-832 <PLU>
A;Cross-references: UNIPROT:Q67892; UNIPARC:UPI00000BF97B; EMBL:Z35716; NID:G527435; PI
A;Experimental source: subtype ayw4
C;Genetics:
A;Gene: P
C;Superfamily: hepatitis virus DNA-directed DNA polymerase
C;Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores: 4.03e-188 Length: 832
Pred. No.: 2651.00 Matches: 494
Score: 95.19% Conservative: 21
Percent Similarity: 91.31% Mismatches: 25
Best Local Similarity: 44.88% Indels: 2
Query Match: 1 Gaps: 0
DB:

US-10-761-006a-1 (1-3215) x S47406 (1-832)

```
QY 1 CTCACAAACATTCACCAAGCTCTGTAGATCCAGGTGAGGGGCTATATTTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerArgSerGlnGlyGluArgProValPheProCys 312
QY 61 TGTGGCTCCAGTTCGCGACAGTAACCTGTGTCGACTACTCTCTCCATATGCTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcysLeuSerHisIleVal 332
QY 121 AATCTTCTCGAGACTGGGGACCTCGACCGAATCGGAGAACCAACATCAGGATTCCT 180
Db 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCTCTGCTGTTTACAGCGGGGTTTCTCTCAATTTCTAGGGGGACCCACGCTTC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGACTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGACCCACGCTTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnTyArgVal-Se 392
QY 300 CTGGCCAAATTCGAGTCCCAACTCCAACTCACTCACCAACTCTGTCTCTCCAAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 412
QY 360 GTCTCTGCTATCGCTGATGTGTCTGCGCGTTTATCATATTTCTCTCTCTCTCTCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGGTCTTCTGACTACCAAGGTATGTCCTGCTCTCTCTCTC 479
Db 432 aMetProHisLeuLeuValGlySerGlyLeuSerArgTyValAlaArgLeuSerSe 452
QY 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATGCAAGACCTGCACACCTCTCTGC 539
Db 452 rAsnSerArgIlePheAsnAsnGlnArgGlyThrMetGlnAsnLeuHisAspTyrcys 472
QY 540 AAGGAACTCTACGTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 472 rArgAsnLeuTyValSerLeuLeuLeuTyGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCTCCATCCATCATCTGCGCTTCGCAGATTCCTATGGAGTGGGCTTCAGTCC 659
Db 492 uTySerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTCTCTCTGGCTCAGTTTACTAGTGCATTTGTTTCTAGTGGTCTTCCGCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
```

```
QY 720 CTGTTTGGCTTTTCAGTTATATGATGATGTTGGGGCGAAGTCTGTGTAACAATCT 779
|
|
|
Db 532 cCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
|
|
|
QY 780 TGAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
|
|
|
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
|
|
|
QY 840 TAAATAAAACCAACGTTGGGCTACTCCCTTAACTTCATGGGATATGTAATTGGGAAGTTG 899
|
|
|
Db 572 oAnlnythrLysAsnGtrpGlyTyrSerLeuAsnPhemetGlyTyrValIleGlyCysTyr 592
|
|
|
QY 900 GGGTACTTTTACCGCGAGCAATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
|
|
|
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
|
|
|
QY 960 TGTAAATAGACCTATTGATGGAAGATATGTCAGAATTTGGGTCTTTTGGGCTTTGCC 1019
|
|
|
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
|
|
|
QY 1020 TGCCCTCTTTACACAATGTTGGCTATCTCTGCTTGTATGCTTTTATATGATGTATACAATC 1079
|
|
|
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuLysProLeuTyrAlaCysIleGlnSe 652
|
|
|
QY 1080 TAAGCAGGCTTTTCACTTTCTCCCAACTTCAAGGCTTTCTGTGTAAACAATATCTGAA 1139
|
|
|
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
|
|
|
QY 1140 CTTTACCCCGTTGCGGCAAGTCCGGTCTCTGCCAAGTGTCTGTGAGCAACCC 1199
|
|
|
Db 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
|
|
|
QY 1200 CACTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGCTCTCT 1259
|
|
|
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
|
|
|
QY 1260 GCCGATCCATCTGCGGAACCTCTAGCAGCTTTGTTTGTCTGCGAGCCGGTCTGAGCAAA 1319
|
|
|
Db 712 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
|
|
|
QY 1320 ACTTATCGAACCGCAACTCTGTGTTCTCTCTCGGAATATACACTCTTCCATGGCT 1379
|
|
|
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 752
|
|
|
QY 1380 GTTAGGCTGTGTCGCAACTGATCTCGCGGACGTCCTTTGTCTAGCTCCCGTCGCGC 1439
|
|
|
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
|
|
|
QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGGCCGTTTGGGGCTCTACCCGTCCTTCCA 1499
|
|
|
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
|
|
|
QY 1500 TCTGCGGTTCCGGCCGACACCGGGCGACCTCTCTTTAGCGGTCTCCCGGTATGTGCC 1559
|
|
|
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
|
|
|
QY 1560 TTTCTATCTGCGGACCGGTGTCACCTTCACCTCTGCGACGTGCGATGGAGACACCC 1619
|
|
|
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
|
|
|
QY 1620 G 1620
|
|
|
Db 832 o 832
```

RESULT 10

JDVLVA

DNA-directed DNA polymerase (BC 2.7.7) - hepatitis B virus (subtype ayw)

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw

C:Date: 18-Dec-1981 #sequence revision 18-Dec-1981 #text_change 09-Jul-2004

C:Accession: A00702; S20748; S53131

R:Gallibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.

Nature 281, 646-650, 1979
A:Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
A:Reference number: A93214; MUID:81012091; PMID:399327
A:Accession: A00702
A:Molecule type: DNA
A:Residues: 1-832 <GAL>
A:Cross-references: UNIPROT:P03156; UNIPARC:UPI00000129801
A:Experimental source: cloned in Escherichia coli
R:Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submitted to the EMBL Data Library, March 1992
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negative
A:Reference number: S20745
A:Accession: S20748
A:Molecule type: DNA
A:Residues: 1-276, 'S', 278-292, 'L', 294-458, 'H', 460-465, 'N', 467-469, 'S', 471-612, 'V', 614-615, 'D', 616-617, 'Q', 618-619, 'M', 620-621, 'L', 622-623, 'S', 624-625, 'D', 626-627, 'S', 628-629, 'L', 630-631, 'S', 632-633, 'D', 634-635, 'S', 636-637, 'L', 638-639, 'S', 640-641, 'D', 642-643, 'S', 644-645, 'L', 646-647, 'S', 648-649, 'L', 650-651, 'S', 652-653, 'L', 654-655, 'S', 656-657, 'L', 658-659, 'S', 660-661, 'D', 662-663, 'S', 664-665, 'L', 666-667, 'S', 668-669, 'L', 670-671, 'S', 672-673, 'L', 674-675, 'S', 676-677, 'L', 678-679, 'S', 680-681, 'D', 682-683, 'S', 684-685, 'L', 686-687, 'S', 688-689, 'L', 690-691, 'S', 692-693, 'L', 694-695, 'S', 696-697, 'L', 698-699, 'S', 700-701, 'D', 702-703, 'S', 704-705, 'L', 706-707, 'S', 708-709, 'L', 710-711, 'S', 712-713, 'L', 714-715, 'S', 716-717, 'L', 718-719, 'S', 720-721, 'D', 722-723, 'S', 724-725, 'L', 726-727, 'S', 728-729, 'L', 730-731, 'S', 732-733, 'L', 734-735, 'S', 736-737, 'L', 738-739, 'S', 740-741, 'D', 742-743, 'S', 744-745, 'L', 746-747, 'S', 748-749, 'L', 750-751, 'S', 752-753, 'L', 754-755, 'S', 756-757, 'L', 758-759, 'S', 760-761, 'D', 762-763, 'S', 764-765, 'L', 766-767, 'S', 768-769, 'L', 770-771, 'S', 772-773, 'L', 774-775, 'S', 776-777, 'L', 778-779, 'S', 780-781, 'D', 782-783, 'S', 784-785, 'L', 786-787, 'S', 788-789, 'L', 790-791, 'S', 792-793, 'L', 794-795, 'S', 796-797, 'L', 798-799, 'S', 800-801, 'D', 802-803, 'S', 804-805, 'L', 806-807, 'S', 808-809, 'L', 810-811, 'S', 812-813, 'L', 814-815, 'S', 816-817, 'L', 818-819, 'S', 820-821, 'D', 822-823, 'S', 824-825, 'L', 826-827, 'S', 828-829, 'L', 830-831, 'S', 832-833, 'L', 834-835, 'S', 836-837, 'L', 838-839, 'S', 840-841, 'D', 842-843, 'S', 844-845, 'L', 846-847, 'S', 848-849, 'L', 850-851, 'S', 852-853, 'L', 854-855, 'S', 856-857, 'L', 858-859, 'S', 860-861, 'D', 862-863, 'S', 864-865, 'L', 866-867, 'S', 868-869, 'L', 870-871, 'S', 872-873, 'L', 874-875, 'S', 876-877, 'L', 878-879, 'S', 880-881, 'D', 882-883, 'S', 884-885, 'L', 886-887, 'S', 888-889, 'L', 890-891, 'S', 892-893, 'L', 894-895, 'S', 896-897, 'L', 898-899, 'S', 900-901, 'D', 902-903, 'S', 904-905, 'L', 906-907, 'S', 908-909, 'L', 910-911, 'S', 912-913, 'L', 914-915, 'S', 916-917, 'L', 918-919, 'S', 920-921, 'D', 922-923, 'S', 924-925, 'L', 926-927, 'S', 928-929, 'L', 930-931, 'S', 932-933, 'L', 934-935, 'S', 936-937, 'L', 938-939, 'S', 940-941, 'D', 942-943, 'S', 944-945, 'L', 946-947, 'S', 948-949, 'L', 950-951, 'S', 952-953, 'L', 954-955, 'S', 956-957, 'L', 958-959, 'S', 960-961, 'D', 962-963, 'S', 964-965, 'L', 966-967, 'S', 968-969, 'L', 970-971, 'S', 972-973, 'L', 974-975, 'S', 976-977, 'L', 978-979, 'S', 980-981, 'D', 982-983, 'S', 984-985, 'L', 986-987, 'S', 988-989, 'L', 990-991, 'S', 992-993, 'L', 994-995, 'S', 996-997, 'L', 998-999, 'S', 1000-1001, 'D', 1002-1003, 'S', 1004-1005, 'L', 1006-1007, 'S', 1008-1009, 'L', 1010-1011, 'S', 1012-1013, 'L', 1014-1015, 'S', 1016-1017, 'L', 1018-1019, 'S', 1020-1021, 'D', 1022-1023, 'S', 1024-1025, 'L', 1026-1027, 'S', 1028-1029, 'L', 1030-1031, 'S', 1032-1033, 'L', 1034-1035, 'S', 1036-1037, 'L', 1038-1039, 'S', 1040-1041, 'D', 1042-1043, 'S', 1044-1045, 'L', 1046-1047, 'S', 1048-1049, 'L', 1050-1051, 'S', 1052-1053, 'L', 1054-1055, 'S', 1056-1057, 'L', 1058-1059, 'S', 1060-1061, 'D', 1062-1063, 'S', 1064-1065, 'L', 1066-1067, 'S', 1068-1069, 'L', 1070-1071, 'S', 1072-1073, 'L', 1074-1075, 'S', 1076-1077, 'L', 1078-1079, 'S', 1080-1081, 'D', 1082-1083, 'S', 1084-1085, 'L', 1086-1087, 'S', 1088-1089, 'L', 1090-1091, 'S', 1092-1093, 'L', 1094-1095, 'S', 1096-1097, 'L', 1098-1099, 'S', 1100-1101, 'D', 1102-1103, 'S', 1104-1105, 'L', 1106-1107, 'S', 1108-1109, 'L', 1110-1111, 'S', 1112-1113, 'L', 1114-1115, 'S', 1116-1117, 'L', 1118-1119, 'S', 1120-1121, 'D', 1122-1123, 'S', 1124-1125, 'L', 1126-1127, 'S', 1128-1129, 'L', 1130-1131, 'S', 1132-1133, 'L', 1134-1135, 'S', 1136-1137, 'L', 1138-1139, 'S', 1140-1141, 'D', 1142-1143, 'S', 1144-1145, 'L', 1146-1147, 'S', 1148-1149, 'L', 1150-1151, 'S', 1152-1153, 'L', 1154-1155, 'S', 1156-1157, 'L', 1158-1159, 'S', 1160-1161, 'D', 1162-1163, 'S', 1164-1165, 'L', 1166-1167, 'S', 1168-1169, 'L', 1170-1171, 'S', 1172-1173, 'L', 1174-1175, 'S', 1176-1177, 'L', 1178-1179, 'S', 1180-1181, 'D', 1182-1183, 'S', 1184-1185, 'L', 1186-1187, 'S', 1188-1189, 'L', 1190-1191, 'S', 1192-1193, 'L', 1194-1195, 'S', 1196-1197, 'L', 1198-1199, 'S', 1200-1201, 'D', 1202-1203, 'S', 1204-1205, 'L', 1206-1207, 'S', 1208-1209, 'L', 1210-1211, 'S', 1212-1213, 'L', 1214-1215, 'S', 1216-1217, 'L', 1218-1219, 'S', 1220-1221, 'D', 1222-1223, 'S', 1224-1225, 'L', 1226-1227, 'S', 1228-1229, 'L', 1230-1231, 'S', 1232-1233, 'L', 1234-1235, 'S', 1236-1237, 'L', 1238-1239, 'S', 1240-1241, 'D', 1242-1243, 'S', 1244-1245, 'L', 1246-1247, 'S', 1248-1249, 'L', 1250-1251, 'S', 1252-1253, 'L', 1254-1255, 'S', 1256-1257, 'L', 1258-1259, 'S', 1260-1261, 'D', 1262-1263, 'S', 1264-1265, 'L', 1266-1267, 'S', 1268-1269, 'L', 1270-1271, 'S', 1272-1273, 'L', 1274-1275, 'S', 1276-1277, 'L', 1278-1279, 'S', 1280-1281, 'D', 1282-1283, 'S', 1284-1285, 'L', 1286-1287, 'S', 1288-1289, 'L', 1290-1291, 'S', 1292-1293, 'L', 1294-1295, 'S', 1296-1297, 'L', 1298-1299, 'S', 1300-1301, 'D', 1302-1303, 'S', 1304-1305, 'L', 1306-1307, 'S', 1308-1309, 'L', 1310-1311, 'S', 1312-1313, 'L', 1314-1315, 'S', 1316-1317, 'L', 1318-1319, 'S', 1320-1321, 'D', 1322-1323, 'S', 1324-1325, 'L', 1326-1327, 'S', 1328-1329, 'L', 1330-1331, 'S', 1332-1333, 'L', 1334-1335, 'S', 1336-1337, 'L', 1338-1339, 'S', 1340-1341, 'D', 1342-1343, 'S', 1344-1345, 'L', 1346-1347, 'S', 1348-1349, 'L', 1350-1351, 'S', 1352-1353, 'L', 1354-1355, 'S', 1356-1357, 'L', 1358-1359, 'S', 1360-1361, 'D', 1362-1363, 'S', 1364-1365, 'L', 1366-1367, 'S', 1368-1369, 'L', 1370-1371, 'S', 1372-1373, 'L', 1374-1375, 'S', 1376-1377, 'L', 1378-1379, 'S', 1380-1381, 'D', 1382-1383, 'S', 1384-1385, 'L', 1386-1387, 'S', 1388-1389, 'L', 1390-1391, 'S', 1392-1393, 'L', 1394-1395, 'S', 1396-1397, 'L', 1398-1399, 'S', 1400-1401, 'D', 1402-1403, 'S', 1404-1405, 'L', 1406-1407, 'S', 1408-1409, 'L', 1410-1411, 'S', 1412-1413, 'L', 1414-1415, 'S', 1416-1417, 'L', 1418-1419, 'S', 1420-1421, 'D', 1422-1423, 'S', 1424-1425, 'L', 1426-1427, 'S', 1428-1429, 'L', 1430-1431, 'S', 1432-1433, 'L', 1434-1435, 'S', 1436-1437, 'L', 1438-1439, 'S', 1440-1441, 'D', 1442-1443, 'S', 1444-1445, 'L', 1446-1447, 'S', 1448-1449, 'L', 1450-1451, 'S', 1452-1453, 'L', 1454-1455, 'S', 1456-1457, 'L', 1458-1459, 'S', 1460-1461, 'D', 1462-1463, 'S', 1464-1465, 'L', 1466-1467, 'S', 1468-1469, 'L', 1470-1471, 'S', 1472-1473, 'L', 1474-1475, 'S', 1476-1477, 'L', 1478-1479, 'S', 1480-1481, 'D', 1482-1483, 'S', 1484-1485, 'L', 1486-1487, 'S', 1488-1489, 'L', 1490-1491, 'S', 1492-1493, 'L', 1494-1495, 'S', 1496-1497, 'L', 1498-1499, 'S', 1500-1501, 'D', 1502-1503, 'S', 1504-1505, 'L', 1506-1507, 'S', 1508-1509, 'L', 1510-1511, 'S', 1512-1513, 'L', 1514-1515, 'S', 1516-1517, 'L', 1518-1519, 'S', 1520-1521, 'D', 1522-1523, 'S', 1524-1525, 'L', 1526-1527, 'S', 1528-1529, 'L', 1530-1531, 'S', 1532-1533, 'L', 1534-1535, 'S', 1536-1537, 'L', 1538-1539, 'S', 1540-1541, 'D', 1542-1543, 'S', 1544-1545, 'L', 1546-1547, 'S', 1548-1549, 'L', 1550-1551, 'S', 1552-1553, 'L', 1554-1555, 'S', 1556-1557, 'L', 1558-1559, 'S', 1560-1561, 'D', 1562-1563, 'S', 1564-1565, 'L', 1566-1567, 'S', 1568-1569, 'L', 1570-1571, 'S', 1572-1573, 'L', 1574-1575, 'S', 1576-1577, 'L', 1578-1579, 'S', 1580-1581, 'D', 1582-1583, 'S', 1584-1585, 'L', 1586-1587, 'S', 1588-1589, 'L', 1590-1591, 'S', 1592-1593, 'L', 1594-1595, 'S', 1596-1597, 'L', 1598-1599, 'S', 1600-1601, 'D', 1602-1603, 'S', 1604-1605, 'L', 1606-1607, 'S', 1608-1609, 'L', 1610-1611, 'S', 1612-1613, 'L', 1614-1615, 'S', 1616-1617, 'L', 1618-1619, 'S', 1620-1621, 'D', 1622-1623, 'S', 1624-1625, 'L', 1626-1627, 'S', 1628-1629, 'L', 1630-1631, 'S', 1632-1633, 'L', 1634-1635, 'S', 1636-1637, 'L', 1638-1639, 'S', 1640-1641, 'D', 1642-1643, 'S', 1644-1645, 'L', 1646-1647, 'S', 1648-1649, 'L', 1650-1651, 'S', 1652-1653, 'L', 1654-1655, 'S', 1656-1657, 'L', 1658-1659, 'S', 1660-1661, 'D', 1662-1663, 'S', 1664-1665, 'L', 1666-1667, 'S', 1668-1669, 'L', 1670-1671, 'S', 1672-1673, 'L', 1674-1675, 'S', 1676-1677, 'L', 1678-1679, 'S', 1680-1681, 'D', 1682-1683, 'S', 1684-1685, 'L', 1686-1687, 'S', 1688-1689, 'L', 1690-1691, 'S', 1692-1693, 'L', 1694-1695, 'S', 1696-1697, 'L', 1698-1699, 'S', 1700-1701, 'D', 1702-1703, 'S', 1704-1705, 'L', 1706-1707, 'S', 1708-1709, 'L', 1710-1711, 'S', 1712-1713, 'L', 1714-1715, 'S', 1716-1717, 'L', 1718-1719, 'S', 1720-1721, 'D', 1722-1723, 'S', 1724-1725, 'L', 1726-1727, 'S', 1728-1729, 'L', 1730-1731, 'S', 1732-1733, 'L', 1734-1735, 'S', 1736-1737, 'L', 1738-1739, 'S', 1740-1741, 'D', 1742-1743, 'S', 1744-1745, 'L', 1746-1747, 'S', 1748-1749, 'L', 1750-1751, 'S', 1752-1753, 'L', 1754-1755, 'S', 1756-1757, 'L', 1758-1759, 'S', 1760-1761, 'D', 1762-1763, 'S', 1764-1765, 'L', 1766-1767, 'S', 1768-1769, 'L', 1770-1771, 'S', 1772-1773, 'L', 1774-1775, 'S', 1776-1777, 'L', 1778-1779, 'S', 1780-1781, 'D', 1782-1783, 'S', 1784-1785, 'L', 1786-1787, 'S', 1788-1789, 'L', 1790-1791, 'S', 1792-1793, 'L', 1794-1795, 'S', 1796-1797, 'L', 1798-1799, 'S', 1800-1801, 'D', 1802-1803, 'S', 1804-1805, 'L', 1806-1807, 'S', 1808-1809, 'L', 1810-1811, 'S', 1812-1813, 'L', 1814-1815, 'S', 1816-1817, 'L', 1818-1819, 'S', 1820-1821, 'D', 1822-1823, 'S', 1824-1825, 'L', 1826-1827, 'S', 1828-1829, 'L', 1830-1831, 'S', 1832-1833, 'L', 1834-1835, 'S', 1836-1837, 'L', 1838-1839, 'S', 1840-1841, 'D', 1842-1843, 'S', 1844-1845, 'L', 1846-1847, 'S', 1848-1849, 'L', 1850-1851, 'S', 1852-1853, 'L', 1854-1855, 'S', 1856-1857, 'L', 1858-1859, 'S', 1860-1861, 'D', 1862-1863, 'S', 1864-1865, 'L', 1866-1867, 'S', 1868-1869, 'L', 1870-1871, 'S', 1872-1873, 'L', 1874-1875, 'S', 1876-1877, 'L', 1878-1879, 'S', 1880-1881, 'D', 1882-1883, 'S', 1884-1885, 'L', 1886-1887, 'S', 1888-1889, 'L', 1890-1891, 'S', 1892-1893, 'L', 1894-1895, 'S', 1896-1897, 'L', 1898-1899, 'S', 1900-1901, 'D', 1902-1903, 'S', 1904-1905, 'L', 1906-1907, 'S', 1908-1909, 'L', 1910-1911, 'S', 1912-1913, 'L', 1914-1915, 'S', 1916-1917, 'L', 1918-1919, 'S', 1920-1921, 'D', 1922-1923, 'S', 1924-1925, 'L', 1926-1927, 'S', 1928-1929, 'L', 1930-1931, 'S', 1932-1933, 'L', 1934-1935, 'S', 1936-1937, 'L', 1938-1939, 'S', 1940-1941, 'D', 1942-1943, 'S', 1944-1945, 'L', 1946-1947, 'S', 1948-1949, 'L', 1950-1951, 'S', 1952-1953, 'L', 1954-1955, 'S', 1956-1957, 'L', 1958-1959, 'S', 1960-1961, 'D', 1962-1963, 'S', 1964-1965, 'L', 1966-1967, 'S', 1968-1969, 'L', 1970-1971, 'S', 1972-1973, 'L', 1974-1975, 'S', 1976-1977, 'L', 1978-1979, 'S', 1980-1981, 'D', 1982-1983, 'S', 1984-1985, 'L', 1986-1987, 'S', 1988-1989, 'L', 1990-1991, 'S', 1992-1993, 'L', 1994-1995, 'S', 1996-1997, 'L', 1998-1999, 'S', 2000-2001, 'D', 2002-2003, 'S', 2004-2005, 'L', 2006-2007, 'S', 2008-2009, 'L', 2010-2011, 'S', 2012-2013, 'L', 2014-2015, 'S', 2016-2017, 'L', 2018-2019, 'S', 2020-2021, 'D', 2022-2023, 'S', 2024-2025, 'L', 2026-2027, 'S', 2028-2029, 'L', 2030-2031, 'S', 2032-2033, 'L', 2034-2035, 'S', 2036-2037, 'L', 2038-2039, 'S', 2040-2041, 'D', 2042-2043, 'S', 2044-2045, 'L', 2046-2047, 'S', 2048-2049, 'L', 2050-2051, 'S', 2052-2053, 'L', 2054-2055, 'S', 2056-2057, 'L', 2058-2059, 'S', 2060-2061, 'D', 2062-2063, 'S', 2064-2065, 'L', 2066-2067, 'S', 2068-2069, 'L', 2070-2071, 'S', 2072-2073, 'L', 2074-2075, 'S', 2076-2077, 'L', 2078-2079, 'S', 2080-2081, 'D', 2082-2083, 'S', 2084-2085, 'L', 2086-2087, 'S', 2088-2089, 'L', 2090-2091, 'S', 2092-209

Db 532 sCysLeuAlaPheSerThrMetAspValLeuGlyAlaLysThrValHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGGTATACATTTAAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnLeuLeuSerLeuGlyLeuHisLeuAsnPr 572
QY 840 TAATAAACCAGAGCTGGGCTACTCCCTTAACCTTCATGGATATGTAATGGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyThrSerLeuHisPheMetGlyThrValIleGlyCys 592
QY 900 GGGTACTTTACCGCAGGAACATATTGACTAAACTCAAGCAATGTTTTCGAAACCTGCC 959
Db 592 rGlySerLeuProGlnAsnHisLeuIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTTGGAAGATGTCGAAAGAAATTTGGGCTTTTGGCTTTC 1019
Db 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACCAATGGCTTATCTGCTGTGATGCTTATGATGCTTATGATGATGATGATC 1079
Db 632 aAlaProPheThrGlnCysGlyThrProAlaLeuMetProLeuThrAlaCysIleGlnSe 652
QY 1080 TAAGCAGCTTTTCACTTTCTCGCAACTTACAGGCTTTCTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrThrLysAlaPheLeuCysLysGlnThrLeuAs 672
QY 1140 CTTTATCCCGTTGCGCGCAACGGTCCGGTCTCTGCAAGTGTGCTGACCAACCC 1199
Db 672 nLeuThrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGATGGGCTTGGCCATAGCCATCAGCGATGCTGGAACCTTTCTGGCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheGlnAlaProLe 712
QY 1260 GCGATCCATACTCGCAACTCTCTAGCAGCTTTGTTGCTCGACGCGGTCTGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACTCTCTTCCATGGCT 1379
Db 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysThrSerPheProTrpLe 752
QY 1380 GTTAGGTGTCTGCCAATCGATCTCTGCGGGAAGTCTTTGTCTACGTCCGCTGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValThrValProSerAl 772
QY 1440 GCTGAATCCCGCGGACGACCTCTCGGCGGCTCTGCGGCTCTACCGTCCCTCTCTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgLysArgLysLeuSerArgProLeuLeuAr 792
QY 1500 TCTGCGGTCTCGGCGGACCGGCGGCACTCTCTTTACGCGGTCTCCCGATGTGCC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuThrAlaAspSerProSerValPr 812
QY 1560 TTCTCATCTCGGACCGTGTGCACTTCTGCTTCACTCTGCACTCGCATCGCATGAGACACC 1619
Db 812 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
QY 1620 G 1620
Db 832 o 832

RESULT 13

S20757

DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype ayw, patient E)

C:Species: hepatitis B virus, HBV

A:Variety: subtype ayw, patient E

C:Date: 20-Feb-1995 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004

C:Accession: S20757

R:Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.

submitted to the EMBL Data Library, March 1992

A:Description: Sequence analysis of HBV genomes isolated from patients with HBSAg negati

A:Reference number: S20745
A:Accession: S20757
A:Molecule type: DNA
A:Residues: 1-832 <LAI>
A:Cross-references: UNIPROT:Q67882; UNIPARC:UPI00000F25B1; EMBL:X65259; NID:959439; PID
A:Experimental source: subtype ayw, patient E
C:Genetics:
A:Gene: P
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 4,38e-187 Length: 832
Score: 2637.00 Matches: 494
Percent Similarity: 95.19% Conservative: 21
Best Local Similarity: 91.31% Mismatches: 25
Query Match: 44.64% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1 (1-3215) x S20757 (1-832)

QY 1 CTCACAACTTCCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTTCTCTGC 60
Db 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGTGGCTCCAGTTCGGGAACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTrpCysLeuSerLeuIleVal 332
QY 121 AATCTTCTCGAGACTGGGGACCTCTCACGAACATGGAGAACACAACTACAGATTCCT 180
Db 333 AsnLeuArgGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352
QY 181 AGGACCTCTCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGAAATCTCAATACC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGGAGCACCACGCTTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgLysAsnTrpArgVal-Se 392
QY 300 CTGGCCAAATTCGCGAGTCCCAACCTCCAATCACTCACCAACTCTGTGCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCTCTGCTATCTCGATGTGTCTGCGCGTTTTATCATATTTCTCTTCTCATCTCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaPheThrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTCTTGTGGTCTTCTGACTACCAAGATATGTTGCCCTTGTCTCTC 479
Db 432 aMetProHisLeuLeuValGlySerGlyLeuSerArgTrpValAlaArgLeuSerSe 452
QY 480 TACTTCCAGGAACATCAACCCAGCAGCGGGCCATGCAAGACTGCGACACTCTCTGCTC 539
Db 452 rAsnSerArgLysPheAsnAsnGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAACTCTACGTTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 472 rArgAsnLeuThrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLe 492
QY 600 GTATTCCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 492 uTrpSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCTGGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTTCTGAGGCTTTTCCCCA 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHI 532
QY 720 CTGTTTGGCTTTTCTAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 532 sCysLeuAlaPheSerThrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552

780 TCAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
| | | | | : : : : :
Db uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572

840 TAAATAAACCAACGTTGGGCTACTCCCTTAACCTCATGGGATATGAATTTGGAAGTTG 899
| | | | | : : : : :
Db oAenLySthrLySArgTrpGlyTyrSerLeuAsnPheMetGlyTyrIleIleGlySerPr 592

900 GCGTACTTTACCGCAGCAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACCTGCC 959
| | | | | : : : : :
Db pGlyThrLeuProGlnAspHisIleValGlnLySleuLySLeuGlyCysPheArgLySLeuPr 612

960 TGTAAATAGACCTATTGATTGAAAGTATGTCAAGAATTTGGGGTCTTTTGGGCTTTGC 1019
| | | | | : : : : :
Db oValAenArgProIleAspTrpLySValTrpGlnArgIleValGlyLeuLeuGlyPheAl 632

1020 TGGCCCTTTTACACAATGGCTATCCCTGCTTGATGCCCTTTATATGATGATATCAATC 1079
| | | | | : : : : :
Db aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl 652

1080 TAAGCAGGCTTTTACCTTTCTCGCAACTTACAGGCTTTTCTGTGTAAACAATATCTGAA 1139
| | | | | : : : : :
Db aLySglnAlaPheThrPheSerProThrTyrLySAlaPheLeuSerLySglnTyrMetAs 672

1140 CCTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCCAAGTGTGCTGACGCCAACCC 1199
| | | | | : : : : :
Db nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692

1200 CACTGATGGGCTTGCCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT 1259
| | | | | : : : : :
Db oThrGlyTyrGlyLeuAlaIleGlyAsnGlnArgMetArgGlyThrIleValAlaProLe 712

1260 GCGGATCCCATACTGCGGAACTCTCTAGCAGCTTGTGTTGCTCGAGCGGCTCGAGCAA 1319
| | | | | : : : : :
Db uProIleHisThrAlaGluLeuAlaIleCysPheAlaArgSerArgSerGlyAlaLy 732

1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATACACCTCTCTTCCATGGCT 1379
| | | | | : : : : :
Db sLeuIleGlyThrAspAsnSerValValLeuSerArgLySThrSerPheProTrpLe 752

1380 GCTAGGGTGTGTCGCAACTGATCTGCGCGGAGCGTCTTTGCTAGTCCGCTCGCGC 1439
| | | | | : : : : :
Db uLeuGlyCysThrAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772

1440 GCTGAATCCCGCGAGCAGCCGCTCGGCGCGTGTGGGGCTCTACCGTCCCTCTTCA 1499
| | | | | : : : : :
Db aLeuAsnProAlaAspProSerArgLyArgLeuGlyLeuSerArgProLeuLeuAr 792

1500 TCTGCGGTTCCGCGCAGCACCGGCGCACCTCTCTTTACGGGTCTCCCGGTATGTGCC 1559
| | | | | : : : : :
Db gLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 812

1560 TTCTCATCTGCGGAGCGGTGTGACCTTTCGCTTTCACCTCTGACGTGCGATGGAGCACCC 1619
| | | | | : : : : :
Db oSerHisLeuProValArgValHisPheAlaSerProLeuHisIleAlaTrpArgProPr 832

1620 G 1620
832 o 832

RESULT 14
S67505
DNA-directed DNA polymerase (BC 2.7.7.7) - hepatitis B virus (subtype ayw3, isolate Hope
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw3, isolate Hope CH1357
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C:Accession: S67505
R:Norde, H.; Ebert, J.W.; Fields, H.A.; Mushahwar, I.K.; Magnus, L.O.
Virology 218, 214-223, 1996
A:Title: Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genob
A:Reference number: S67503; MUID:96207410; PMID:8615024
A:Accession: S67505
A:Molecule type: DNA

A:Residues: 1-832 <NOR>
A:Cross-references: UNIPROT:P87744; UNIPARC:UPI0000074B8; EMBL:U46935; NID:g1814218; PI
A:Experimental source: subtype ayw3, isolate Hope CH1357
A>Note: the authors translated the codon GAA for residue 290 as Ala
C:Genetics:
A:Gene: P
A:Introns: 292/3
C:Superfamily: hepatitis virus DNA-directed DNA polymerase
C:Keywords: DNA biosynthesis; nucleotidyltransferase

Alignment Scores:
Pred. No.: 2,03e-186 Length: 832
Score: 2628.00 Matches: 490
Percent Similarity: 95.1% Conservative: 25
Best Local Similarity: 90.5% Mismatches: 25
Query Match: 44.4% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x S67505 (1-832)

QY 1 CTCCACAACATTTCCACCAAGCTCTCTAGATCCCAAGGTGAGGGGCCCTATATTTTCCTGC 60
| | | | | : : : : :
Db 293 LeuTyrSerIleProProAsnSerAlaArgSerGlnSerThrGlyProIleLeuSerCys 312

61 TGGTGGCTCCAGTTCCTCGGAACAGTAAACCTGTGTTCCGACTACTGCCTCTCCCATATCGTC 120
| | | | | : : : : :
Db 313 TrpTrpLeuGlnPheArgAsnSerGluProCysSerAspTyr-CysLeuSerHisLeuVal 332

121 AATCTTCTCGAGGACTGGGGACCTCGACCGAACAATGAGAGACACACATCAGGATTCCT 180
| | | | | : : : : :
Db 333 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 352

181 AGGACCCCTGCTGCTGTACAGCGGGGTTTTCTCGTTGACAAGAACTCTCACAAATACC 240
| | | | | : : : : :
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLyAsnProHisAsnThr 372

241 CGAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTTAGGGGAGCACCCACGTGTTCT 299
| | | | | : : : : :
Db 373 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 392

300 CTGGGCAAAATTCGAGTCCCCCAACTCCAATCCTCACCACTCTGTCTCTCCAAATTT 359
| | | | | : : : : :
Db 392 rTrpProLySPhaAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412

360 GTCTTGGCTATCGCTGGATGTCTGCGCGGTTTTATCATATTCCTTCTCATCTCTGCTGC 419
| | | | | : : : : :
Db 412 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432

420 TATGCTCATCTTCTTGTGTTGTTCTTCTGGACTTACCAAGGTATGTTGCCGTTTGTCTTC 479
| | | | | : : : : :
Db 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452

480 TACTTCCAGGAACATCAACACGAGCCGAGCCATGCAAGACCTGCAGACTCTCTGCTC 539
| | | | | : : : : :
Db 452 rThrSerArgIleIleAspHisGlnHisGlyThrMetGlnAsnLeuHisAspHisCysSe 472

540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAAACTGCACATT 599
| | | | | : : : : :
Db 472 rArgAsnLeuPheValSerLeuLeuLeuTyrLySThrPheGlyArgLySLeuHisLe 492

600 GTATTCCCATCCCATCATCTCTGGGCTTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
| | | | | : : : : :
Db 492 uTyrSerHisProIleValLeuGlyPheArgLySleIleProMetGlyValGlyLeuSerPr 512

660 GTTTCCTCTGGCTCAGTTTACTAGTCCCATTTGTTGAGTGGTTCGTAGGGCTTCCCCCA 719
| | | | | : : : : :
Db 512 oPheLeuLeuAlaGlnPheThrSerSerIleCysSerValValArgArgAlaPheProHi 532

720 CTGTTTGGCTTTTCAGTTATATGATGATGTTGTTGTTGGGGCGAAGCTCTGTACCAATCT 779
| | | | | : : : : :
Db 532 sCysLeuAlaPheSerTyrMetAspAspLeuValLeuGlyAlaLySerValGlnHisLe 552

780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
| | | | | : : : : :
QY

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:40:32 ; Search time 180.238 Seconds
(without alignments)
3320.243 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GTCTTTGGGTATACATTAA 681

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DBV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10761006/runat_27122005_192900_15743/app_query.fasta_1.1102
-DB=A-Geneseq -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPT=0 -LOOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=delop62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006@cgn 1.1 476 @runat_27122005_192900_15743 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	96.2	400	3	AAY54045 Amino aci
2	1206	94.4	226	1	AAP40310 Hepatitis
3	1206	94.4	226	3	AAB35768 Hepatitis
4	1206	94.4	226	4	AAB98043 Synthetic
5	1206	94.4	226	4	AAB80958 Viral pro
6	1206	94.4	226	5	ABB77789 Amino aci
7	1206	94.4	226	5	ABB05220 Hepatitis
8	1206	94.4	226	5	AAB97647 Hepatitis
9	1206	94.4	226	8	ADH77879 Hepatitis

10	1206	94.4	318	2	AAR60174
11	1206	94.4	383	1	AAP60416
12	1206	94.4	406	8	ADH77843 Hepatitis
13	1206	94.4	406	8	ADO48159 Hepatitis
14	1206	94.4	406	8	AD91504 Amino aci
15	1206	94.4	406	9	AEA21247 Nanoparti
16	1206	94.4	582	7	ADD69670 Chimeric
17	1206	94.4	590	7	ADD69668 Chimeric
18	1206	94.4	658	7	ADD69666 Chimeric
19	1206	94.4	1109	7	ADD69672 Chimeric
20	1203	94.1	226	5	ABR05221 Hepatitis
21	1203	94.1	226	5	AU97648 Hepatitis
22	1203	94.1	236	1	AAP81992 Hepatitis
23	1200	93.9	226	5	AAW49180 HBV subty
24	1200	93.9	396	2	AAR60177 Chimeric
25	1199	93.8	281	1	AAP60163 Subtype a
26	1199	93.8	281	1	AAP60560 Hepatitis
27	1199	93.8	281	1	AAP60617 C-termina
28	1199	93.8	385	2	AAR60176 Chimeric
29	1198	93.7	226	9	AEA08836 Hepatitis
30	1197	93.7	400	2	AAU93801 Hepatitis
31	1197	93.7	400	6	ABR55863 HBV S-pro
32	1197	93.7	400	8	ADM49718 Hepatitis
33	1197	93.7	400	8	ADU74394 HBV S-pro
34	1195	93.5	281	1	AAP70294 Subtype a
35	1195	93.5	499	2	AAR60175 Chimeric
36	1193	93.3	400	4	AAG66930 HBV genot
37	1192	93.3	344	9	AEA08851 Hepatitis
38	1190	93.1	269	4	AAW48418 Protein #
39	1190	93.1	281	2	AAR62870 Hepatitis
40	1190	93.1	281	4	AAW48419 Protein #
41	1190	93.1	389	1	AAP60794 Atr-type
42	1190	93.1	843	3	AAY54044 Amino aci
43	1189	93.0	226	3	AAU93739 A wild ty
44	1188	93.0	226	3	AAW48415 HBV surfa
45	1187	92.9	226	9	AEA08848 Hepatitis

ALIGNMENTS

RESULT 1
ID AAY54045 standard; protein; 400 AA.
XX AC AAY54045;
XX AC AAY54045;
XX 27-MAR-2000 (first entry)
XX Amino acid sequence of a HBV large surface antigen protein.
XX HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase;
XX large surface antigen; core protein; transactivating X protein;
XX hepatitis vaccine; HBV infection; hepatocellular carcinoma.
XX Hepatitis B virus.

XX Key Location/Qualifiers
XX Misc-difference 14 /note= "encoded by AGC"
XX Misc-difference 174..400 /note= "these residues are specifically claimed in claim 29"
XX Misc-difference 298..320 /note= "these residues are specifically claimed in claim 23"

XX WO9966048-A1.
XX 23-DEC-1999.
XX 19-JUN-1998; 98WO-SG000046.
XX 19-JUN-1998; 98WO-SG000046.

```

XX PA (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.
XX PI
XX PI Oon CJ, Lim GK, Zhao Y, Chen WN;
XX XX
XX DR WPI; 2000-106104/09.
XX DR N-PSDB; AA237088.
XX PT New isolated hepatitis B virus strain, useful for, e.g. treatment of
XX PT hepatitis infection.
XX PS Claim 23; Page 39-40; 68pp; English.
XX XX
CC The present sequence is encoded by the genome of an isolated strain of
CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
CC antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited as
CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
CC large surface antigen, a core protein, and a transactivating X protein.
CC The large surface antigen differs from the wild type sequence in that it
CC contains a Thr at position 133 of the wild type sequence instead of a
CC Met. The proteins are used to produce antibodies. The proteins,
CC polynucleotide and antibodies can be used for detecting the novel HBV
CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
CC HBV novel strain polypeptides can be used to identify compounds for
CC treating or preventing HBV infection or hepatocellular carcinoma
XX XX
SQ Sequence 400 AA;

Alignment Scores:
Pred. No.: 4,35e-112 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 3 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAY54045 (1-400)

QY 1 ATGGAGAACACATCAGATTCTAGGACCCCTGCTGCTTACAGGGGGTTTTTC 60
Db 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
Db 195 SerLeuThrArgLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTTCTAGGGGAGACCCACGTTCTCGCCAAATTCGACGTCGCCAACCTCCAATCA 179
Db 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCAACTCTTGTCTCCCAATTTGTTCTCGCTATCCTCGATGCTGTCGGGGTTT 239
Db 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpAsnCysLeuArgArgPh 254
QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTGTTGTTCTTCTGACTA 299
Db 254 elleillePheLeuPheIleleuleuLeuCysleullePheLeuLeuValleuAspTy 274
QY 300 CCAAGTATGTGCGGTTGCTTGTCTTACTTTCAGGAACATCAACACCAGCACGGGGCC 359
Db 274 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGCAGAGCTGCACACACTCTGCTCAAGGAACCTAGCTTCCCTCTGTTGCTGTAC 419
Db 294 oCyslysthrCysThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysTh 314
QY 420 AAAACCTTCGGACGGAAACTGCACTTGTATTCCATCCCATCCATCATCTCGGGCTTTCGAAG 479
Db 314 rlysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 334
QY 480 ATTCTATGGAGTGGGGCTCAGTCCGTTTCTCTCGGCTCAGTTTACTAGTGCCATTGTT 539

```

```

Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 354
QY 540 TCAGTGGTTCGTAGGGCTTCCCCACTGTTTGGCTTTTCAGTTATATGATGATGTTA 599
Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGCGAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
Db 374 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyIle 400

RESULT 2
ID AAP40310 standard; protein; 226 AA.
XX AAP40310;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 18-AUG-1992 (first entry)
XX
DE Hepatitis virus adr-B surface antigen.
XX
KW HBsAg; vaccine; diagnosis; HBV infection.
XX
OS Hepatitis B virus.
XX
PN JP59074985-A.
XX
PD 27-APR-1984.
XX
PF 19-OCT-1982; 82JP-00183432.
XX
PR 19-OCT-1982; 82JP-00183432.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
WPI; 1984-143231/23.
XX
DNA used in prevention of infections by hepatitis virus B - comprises
PT structural gene of hepatitis virus adr-B surface antigen, coding gene and
PT at least 1 virus core antigen structural gene.
XX
PS Disclosure; Fig 2; 13pp; Japanese.
XX
CC The sequence is that of hepatitis virus adr-B surface antigen (HBsAg). It
CC can be used as a vaccine for the prevention of infections by hepatitis B
CC virus (HBV). See also AAP40311. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8,82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAP40310 (1-226)

QY 1 ATGGAGAACACATCAGATTCTAGGACCCCTGCTGCTTACAGGGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40

```

QY 120 TTTCTAGGGGACACCCAGCTGTTCTGCGCAAAATTCGAGTCCCAACCTCCCAATCA 179
 Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTGTCTCCCAATTTGTCTGCTATCGCTGATGCTGCGGGGTTT 239
 Db 60 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTCCTCTTCATCCTGCTGCTATGCTCATCTTCTTCTGTTCTTCTGACTA 299
 Db 80 eillePheLeuPheLeuLeuLeuLeuCysLeuLeuPheLeuValLeuLeuAspTy 100
 QY 300 CCAAGTATGTTCCCGTTGCTCTCTACTTCCAGGAACATCAACACGACGCGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
 QY 360 ATGCAAGACTGTCAGCACTCTCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
 Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGAGCGGAACTGCTATTCATCCCATCCCATCTCTGGGCTTTCGCAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAla 160
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTGCTCAGTTACTAGTGCATTGT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGTTCGTAGGGCTTCCCGCACTGTTTCGCTTTCAGTTATATGGATGATGGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 200
 QY 600 TTGGGGCGGAACTGTGACAACTCTTGAGTCCCTTTTACCTCTAFTACCAATTTCTT 659
 Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuProllePhePh 220
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValTyIle 226
 RESULT 3
 AAB35768
 ID AAB35768 standard; protein; 226 AA.
 AC
 XX AAB35768;
 DT 06-AUG-2003 (revised)
 DT 21-FEB-2001 (first entry)
 XX
 DE Hepatitis b virus protein sequence SEQ ID 12.
 XX
 KW Viral antigen preparation; peptidyl prolyl cis-trans isomerase; PPI;
 KW immune diagnosis.
 XX
 OS Hepatitis B virus.
 XX
 PN JP2000249704-A.
 XX
 PD 14-SEP-2000.
 XX
 PF 17-MAY-1999; 99JP-00136346.
 XX
 PR 28-DEC-1998; 98JP-00377103.
 XX
 PA (SEKI) SEKISUI CHEM IND CO LTD.
 PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
 XX
 DR WPI; 2000-622348/60.
 DR N-PSDB; AAC66342.
 XX
 PT Virus antigen preparation for immune diagnosis involves processing virus
 PT antigen in the presence of peptidyl prolyl cis-trans isomerase.

XX Example 7; Page 19-20; 26pp; Japanese.
 PS This invention relates to a method for the preparation of a viral
 CC antigen. The method involves processing the virus antigen in the presence
 CC of a peptidyl prolyl cis-trans isomerase (PPI). Preferably a surfactant,
 CC a reducer and a protein denaturant are also added. The method is used for
 CC immune diagnosis. The present sequence represents a Hepatitis b virus
 CC protein sequence used in the invention. (Updated on 06-AUG-2003 to
 CC correct OS field.)
 XX
 SQ Sequence 226 AA;
 Alignment Scores: Length: 226
 Pred. No.: 8,826-110 Matches: 220
 Score: 1206.00
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 3 Gaps: 0
 US-10-761-006A-1_COPY_155_835 (1-681) x AAB35768 (1-226)
 QY 1 ATGAGAACACCAATCAGATTCTAGGACCCCTCTGCTGTATACAGCGGGGTTTTC 60
 Db 1 MetGlnAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePh 20
 QY 61 TCGTTGACAAGAATCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGAGCAGCCAGCTGTTCTGCGCAAAATTCGAGTCCCAACCTCCCAATCA 179
 Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTGTCTCCCAATTTGTCTGCTATCGCTGATGCTGCGGGGTTT 239
 Db 60 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTCCTCTTCATCCTGCTGCTATGCTCATCTTCTTCTGTTCTTCTGACTA 299
 Db 80 eillePheLeuPheLeuLeuLeuCysLeuLeuPheLeuValLeuLeuAspTy 100
 QY 300 CCAAGTATGTTCCCGTTGCTCTCTACTTCCAGGAACATCAACACGACGCGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
 QY 360 ATGCAAGACTGTCAGCACTCTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
 Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGAGCGGAACTGCTATTCATCCCATCCCATCTCTGGGCTTTCGCAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAla 160
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTGCTCAGTTACTAGTGCATTGT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGTTCGTAGGGCTTCCCGCACTGTTTCGCTTTCAGTTATATGGATGATGGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 200
 QY 600 TTGGGGCGGAACTGTGACAACTCTTGAGTCCCTTTTACCTCTAFTACCAATTTCTT 659
 Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuProllePhePh 220
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValTyIle 226
 RESULT 4
 AAB98043

```
ID XX AAB98043 standard; protein; 226 AA.
XX AC AAB98043;
XX DT 15-AUG-2001 (first entry)
XX DE Synthetic adr type hepatitis B surface antigen SEQ ID NO:12.
XX KW Methanococcus; Hepatitis B virus; Hepatitis C virus; HIV; PPIase;
XX KW hepatitis B surface antigen; hepatitis B core antigen; virus antigen;
XX KW peptidyl prolyl cis-trans isomerase; virucide; antiinflammatory;
XX KW anti-HIV; anti-hepatitis B; antihepatitis-C; hepatotropic; vaccine;
XX KW immunodiagnostic.
OS Hepatitis B virus.
OS Synthetic.
XX PN JP2001066309-A.
XX PD 16-MAR-2001.
XX PF 27-SEP-1999; 99JP-00273203.
XX PR 28-DEC-1998; 98JP-00377103.
XX PR 17-MAY-1999; 99JP-00136346.
XX PA (SEKI ) SEKISUI CHEM IND CO LTD.
XX PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX DR WPI; 2001-304300/32.
XX DR N-PSDB; AAH21782.
XX PT Preparation of viral antigen, especially hepatitis B, hepatitis C or HIV,
XX PT for use as vaccine and in diagnostics, comprises reacting peptidyl prolyl
XX PT cis-trans isomerase to the viral antigen.
XX PS Example 7; Page 21-22; 31pp; Japanese.
XX CC The present invention describes a method (M) for the preparation of a
XX CC viral antigen (I) comprising reacting (I) with peptidyl prolyl cis-trans
XX CC isomerase (PPIase). (I) has virucide, antiinflammatory, anti-HIV, anti-
XX CC hepatitis B, antih hepatitis-C and hepatotropic activities. (I) is useful
XX CC for immunodiagnosics, and as a vaccine. (I) has stronger antigenic
XX CC activity than previously available antigens. Large quantities of viral
XX CC antigen can be prepared by the simple process within a short culture
XX CC period. The present sequence represents a synthetic adr type hepatitis B
XX CC surface antigen, which is given in an example from the present invention
XX SQ Sequence 226 AA;

Alignment Scores:
Pred. No.: 8,82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 4 Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x AAB98043 (1-226)
QY 1 ATGGAGAACACACATCAGATTCTCTAGGACCCCTGCTGTTACAGCGGGGTTTTC 60
Db 1 MetGluasnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTCACAAAGATCCTCAATACCGCAGAGTCTAGATC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgileLeuThrileProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTCTAGGGGGAGACCCACAGTGTCTCGSCCAAAATTCGACAGTCCCAACTCCCAATCA 179
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH 60
QY 180 CTCACCAACCTCTTGTCTCTCAATTGTCTCGCTATCGTGTGATGTGTCTGGCGGTTT 239
```

```
|||||
60 sSerProThrSerCysProProlleCysProGlyTyzArgTrpMetCysLeuArgArgPh 80
240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTGTTGGTCTTCTGGACTA 299
80 eillePheLeuPheilleLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspIy 100
300 CCAAGGTATGTGCGGTTTGTCTCTACTTCCAGGAACATCAACCAACGACGCGGCC 359
100 xGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
360 ATGCAAGACCTGACGACTCTCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGTGTGAC 419
120 oCysLysThrCysThrilleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
420 AAAACCTTCGACGGAACCTGCTGTTATTCCTCCATCCATCCCTGGCTTTCGCAAG 479
140 rLysProSerAspGlyAsnCysThrCysilleProleProSerSerTrpAlaPheAlaR 160
480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTCCATTGT 539
160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
540 TCAGTGTTCGTAGGGCTTCCCCACATGTTGGCTTTTTCAGTTATATGATGATGTGTA 599
180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValilleTrpMetMetTrpTy 200
600 TTGGGGGCGAAGTCTGTACACATCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
200 rTrpGlyProSerLeuTyfAsnilleLeuSerProPheLeuProleLeuPhePh 220
660 TTGTCTTTGGGTATACATT 678
220 eCysLeuTrpValTyrlle 226

RESULT 5
AAB80958
ID AAB80958 standard; protein; 226 AA.
XX AC AAB80958;
XX DT 08-JUN-2001 (first entry)
XX DE Viral protein sequence #4.
XX KW Virus antigen; molecular chaperone; immunoassay reagent.
XX OS Unidentified.
XX PN JP2001033449-A.
XX PD 09-FEB-2001.
XX PF 27-SEP-1999; 99JP-00273204.
XX PR 28-DEC-1998; 98JP-00377102.
XX PR 17-MAY-1999; 99JP-00136334.
XX PA (SEKI ) SEKISUI CHEM IND CO LTD.
XX PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX DR WPI; 2001-230268/24.
XX DR N-PSDB; AAF81488.
XX PT Preparation of a virus antigen.
XX PS Example 7; Page 23; 33pp; Japanese.
XX CC The present invention relates to a method for the preparation of a virus
XX CC antigen. The method consists of a step of reacting a molecular chaperone
XX CC with a virus antigen. The virus antigen is useful for the preparation of
XX CC immunoassay reagents. The present sequence is a viral protein, which was
XX CC used in the present invention
```

```

XX SQ Sequence 226 AA;
Alignment Scores:
Pred. No.: 8.82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 4 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAB80958 (1-226)
QY 1 ATGAGAGACACATCAGATTCCTAGGACCCCTGCTGTGTACAGGGGGGTTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACATACGACGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGCACCACGCTGTTCTGGCCAAATTCGACGTCCCAACCTCCCAATCA 179
DB 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGTCTCTCAATTTGCTCGCTATCGCTGGANGTCTGGCGGTTT 239
DB 60 sSerProThrSerCysProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCTCTTCATCTCTGCTGTATGCTCCATCTTCTGTGTCTTCTTGGACTA 299
DB 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTrp 100
QY 300 CCAAGGTATGTTGCCGCTTGTCTCTTCTTCTCTTCTGAGGAACTTCCAGGAACTCA 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGCACGACTCTCTCTCAAGAAACTCTTACCTTCCCTCTTGTGTCTGTAC 419
DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGACTGTATTCATCCATCCCATCATCTCTGGGCTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTCTATGGAGTGGGCTTCAGTCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180

RESULT 6
ID ABB77789
XX ABB77789 standard; protein; 226 AA.
AC ABB77789;
XX
XX 20-AUG-2002 (first entry)
XX
XX Amino acid sequence of a hepatitis B virus (HBV) protein.
XX HBV; surface antigen; HBs antigen.
XX

```

```

OS Hepatitis B virus.
XX JP2002101887-A.
XX 09-APR-2002.
XX 29-SEP-2000; 2000JP-00299262.
XX 29-SEP-2000; 2000JP-00299262.
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX WPI: 2002-440457/47.
XX N-PSDB; ABL59070.
XX Preparing a recombinant hepatitis B virus surface antigen and an antigen.
XX Disclosure; Page 8-9; 9pp; Japanese.
XX The specification describes a method for preparing a recombinant
XX hepatitis B virus (HBV) surface (HBs) antigen. The recombinant HBs
XX antigen is useful in the preparation of immunoassay reagents. The present
XX sequence represents a HBV protein, which was used in the course of the
XX invention
XX SQ Sequence 226 AA;
Alignment Scores:
Pred. No.: 8.82e-110 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 4 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x ABB77789 (1-226)
QY 1 ATGAGAGACACATCAGATTCCTAGGACCCCTGCTGTGTACAGGGGGGTTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACATACGACGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGCACCACGCTGTTCTGGCCAAATTCGACGTCCCAACCTCCCAATCA 179
DB 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTTGTCTCTCAATTTGCTCGCTATCGCTGGANGTCTGGCGGTTT 239
DB 60 sSerProThrSerCysProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCTCTTCATCTCTGCTGTATGCTCCATCTTCTGTGTCTTCTTGGACTA 299
DB 80 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTrp 100
QY 300 CCAAGGTATGTTGCCGCTTGTCTCTTCTCTTCTCTTCTGAGGAACTTCCAGGAACTCA 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGCACGACTCTCTCTCAAGAAACTCTTACCTTCCCTCTTGTGTCTGTAC 419
DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGACTGTATTCATCCATCCCATCATCTCTGGGCTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTCTATGGAGTGGGCTTCAGTCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180

```


CC containing four antigen sites of the S antigen of the hepatitis B virus
 CC surface antigen and a region encoding at least four amino acids forming
 CC the membrane-penetrating region. The invention also comprises a hepatitis
 CC B virus surface antigen prepared using this method. The method of the
 CC invention is useful for the preparation of immunoassay reagents. The
 CC present sequence represents the Hepatitis B surface antigen protein #3 of
 CC the invention

XX SQ Sequence 226 AA;

Alignment Scores:
 Pred. No.: 8,82e-110 Length: 226
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 5 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAU97647 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTCTGTTACAGCGGGGTTTTTC 60
 Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCCTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGGAGCACCACAGTGCTCTGCGCCAAAATTCGAGTCCCACTCCCAATCA 179
 Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTTGCTCCCAATTTGCTGCTATCGTGGTATCGTGGTATCGTGGGTTT 239
 Db 60 sSerProThrSerCysProPheLeuCysProGlyTyrArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTCCTCTTCATCTCTGCTATGCTGCTCATCTCTTCTTGTGTTCTTCTGACTA 299
 Db 80 elIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGTATGTTGCCGTTTGTCTCTTACTTTCAGGAACATCAACACAGCAGCGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGCTGTAC 419
 Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGAGCGGAACGCACTTGATATCCATCCCATCATCTCGGGCTTCGCAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
 QY 480 ATTCTATGGAGTGGGCGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCATTGT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
 QY 540 TCAGTGTGTTAGGGCTTTCCCGCACTGTTTGGCTTTTCAAGTATATGAGTATGGGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
 QY 600 TTGGGGCGGAGTCTGTACACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 200 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValIle 226

RESULT 9

ADH77879

ID ADH77879 standard; protein; 226 AA.

XX

AC ADH77879;

XX 22-APR-2004 (first entry)
 XX Hepatitis B virus surface antigen, HBsAgS.
 XX Nanoparticle; surface antigen protein; HBsAgS; liver; drug delivery.
 XX Hepatitis B virus.
 XX WO2004002459-A1.
 XX 08-JAN-2004.
 XX 27-JUN-2003; 2003WO-JP008244.
 XX 28-JUN-2002; 2002JP-00191386.
 XX 27-JUN-2003; 2003JP-00183863.
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX Kuroda S, Tanizawa K, Kondo A, Ueda M, Seno M, Tada H;
 XX WPI; 2004-082876/08.
 XX Hollow nanoparticle comprising protein with modified cysteine residues
 XX for target specific drug delivery.
 XX Example 1; Fig 2; 79pp; Japanese.
 XX The present invention relates to a hollow nanoparticle, which comprises a
 XX protein, e.g. hepatitis B virus surface antigen protein, with modified
 XX cysteine residues. The nanoparticle specifically recognizes a target cell
 XX e.g. a liver cell. The nanoparticle is useful in methods for drug
 XX delivery. The present sequence is one such hepatitis B virus surface
 XX antigen protein.
 XX SQ Sequence 226 AA;

Alignment Scores:

Pred. No.: 8,82e-110 Length: 226
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 5 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x ADH77879 (1-226)

QY 1 ATGGAGAACACATCAGGATTCCTAGGACCCCTGCTCTGTTACAGCGGGGTTTTTC 60
 Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCCTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGGAGCACCACAGTGTTCTGCGCCAAAATTCGAGTCCCACTCCCAATCA 179
 Db 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTTGCTCCCAATTTGCTGCTATCGTGGTATCGTGGTATCGTGGGTTT 239
 Db 60 sSerProThrSerCysProPheLeuCysProGlyTyrArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTCCTCTTCATCTCTGCTATGCTGCTCATCTCTTCTTGTGTTCTTCTGACTA 299
 Db 80 elIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGTATGTTGCCGTTTGTCTCTTACTTTCAGGAACATCAACACAGCAGCGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGCTGTAC 419

Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGACATCTGATTCCTCCATCCCATCATCTGGGCTTTGGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGGAGTGGGCTCAGTCGCTTCTCTGGCTCAGTTTACTAGTGCCTTTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTTCGTAGGCTTCCGCCACCTGTTGGCTTTTCAGTATATGATGATGCTA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuProIlePhePh 220
QY 660 TTGTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 10

AAR60174
ID AAR60174 standard; protein; 318 AA.
XX
AC AAR60174;
XX
DT 16-OCT-2003 (revised)
DT 07-APR-1995 (first entry)
XX
Chimeric protein (R3:S) used in multivalent vaccine.
XX
Chimeric; chimera; vaccine; multivalent; hepatitis B virus; HBV;
KW hepatitis; Japanese encephalitis virus; baculovirus.
XX
OS Hepatitis B virus.
OS Japanese encephalitis virus.
OS Chimeric.
XX JP06205672-A.
XX 26-JUL-1994.
XX
XX 19-MAR-1992; 92JP-00063699.
XX 19-MAR-1992; 92JP-00063699.
XX (JAPG) NIPPON ZEON KK.
PA (TOKS-) TOKYO SHINKAI KAGAKU SOGO KENKYUSHO ZH.
XX
XX WPI; 1994-275516/34.
DR N-PSDB; AAQ70155.
XX
PT Prodn. of chimeric proteins having antigenic sites from Japanese
PT encephalitis virus and hepatitis B virus surface antigens - also
PT recombinant baculovirus, useful as multivalent vaccine.
XX
PS Disclosure; Fig 1-3; 13pp; Japanese.
XX
CC The chimeric protein comprises antigenic sites from Japanese encephalitis
CC virus and Hepatitis B virus surface antigens. The protein may be used as
CC a multivalent vaccine. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 318 AA;

Alignment Scores:
Pred. No.: 9.55e-110 Length: 318
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2

DB: 2 Gaps: 0
US-10-761-006A-1_COPY_155_835 (1-681) x AAR60174 (1-318)
QY 1 ATGGAGAACACAAATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
Db 93 MetGluAsnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 112
QY 61 TCGTTGACAAAGATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGAGCTTCTCTCAAT 119
Db 113 LeuLeuThrAglIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 132
QY 120 TTTCTAGGGGAGACCCACGCTGTTCTCGGCCAAAAATTCGCAGTCCCCAACCTCCAATCA 179
Db 133 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 152
QY 180 CTCACCAACCTTGTCTCCAAATTTGTCCTGGCTATCGCTGGATGTCTGCGGCGTTT 239
Db 152 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 172
QY 240 TATCATATTCCTTTTCATCTGCTATGCTCATCTTCTTTGGTTCCTTCGACTA 299
Db 172 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 192
QY 300 CCAAGGTATGTGCCCGTTTGTCTTACTTCCAGGAAACATCAACACAGCAGCGGGCC 359
Db 192 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 212
QY 360 ATGCAAGACCTGCGAGCTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTTCGTGTAC 419
Db 212 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 232
QY 420 AAAACCTTCGACGGAACACTGCACCTTGATTTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 232 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 252
QY 480 ATTCTATGGAGTGGGCTCAGTCCGCTTCTCCGGCTCAGTTTACTAGTGCCTTTGT 539
Db 252 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 272
QY 540 TCAGTGTTCGTAGGCTTCCGCCACCTGTTGGCTTTTCAGTATATGATGATGCTA 599
Db 272 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 292
QY 600 TTGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 292 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuPhePh 312
QY 660 TTGTCTTTGGGTATACATT 678
Db 312 eCysLeuTrpValTyIle 318

RESULT 11

AAP80416
ID AAP80416 standard; protein; 383 AA.
XX
AC AAP80416;
XX
DT 25-MAR-2003 (revised)
DT 12-SEP-1990 (first entry)
XX
DE Sequence of adr type hepatitis B virus (HBV) surface antigen (HBeAg) L
DE protein (M protein, S protein).
XX
KW Egg white lysozyme; hepatitis B virus (HBV) surface antigen (HBeAg);
KW Saccharomyces cerevisiae AH22R/PGLD LP39-Rct; L protein; M protein;
KW Saccharomyces cerevisiae LILP39-Rct; Saccharomyces cerevisiae LP31-Rct;
KW S protein.
OS Hepatitis B virus.
XX
PN EP288198-A.
XX

PD 26-OCT-1988.
 XX
 PF 13-APR-1988; 88EP-00303297.
 XX
 PR 20-APR-1987; 87JP-00098265.
 PR 12-OCT-1987; 87JP-00256885.
 PR 18-APR-1988; 88JP-00095335.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Fujisawa Y, Imai S, Miyazaki T;
 XX
 DR WPI; 1988-301233/43.
 DR N-PSDB; AAN80973.
 XX
 PT Peptide(s) having hepatitis B surface antigenicity - used as vaccine for
 PT prevention of hepatitis B virus infection and in diagnostic kits.
 XX
 PS Example; Fig 5; 31pp; English.
 XX
 CC HBsAg L protein (P39) is an env protein. When L protein genes are
 CC introduced into animal cells (CHO cells) they produce HBsAg particles
 CC containing M and S proteins. The advantage is that peptides having HBsAg
 CC antigenicity are excreted outside the cells and easily purified. Claimed
 CC is a eukaryotic cell transformed with the rDNA. Pref. the cell is a yeast
 CC cell, esp. S. cerevisiae AH22R-/pGLD LP39- Rct, LIP39-Rct or LP31-Rct.
 CC Also claimed is recombinant DNA, which is a DNA coding for a signal
 CC peptide which functions in a eukaryotic cell, and is bound to the 5'-
 CC terminal of a DNA coding for a peptide having HBsAg activity. The signal
 CC peptide may be a signal peptide of egg white lysozyme. When trypsin-like
 CC protease-producing yeast is utilized as the host, L protein and M protein
 CC may possibly be degraded by the protease; therefore it is desirable that
 CC the genes are altered so that the 48th arginine residue from the N-
 CC terminal of M protein or a peptide containing the residue (preferably the
 CC 44-49 peptide) may be deleted. Preferable genes include L protein gene
 CC (altered) coding for the amino acid sequence 1-383 shown in AAN80973.
 CC (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
 CC correct PA field.)
 XX
 SQ Sequence 383 AA;

Alignment Scores:
 Pred. No.: 9.98e-110 Length: 383
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x AAP80416 (1-383)

QY 1 ATGGAGAACACAACTCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTC 60
 Db 158 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 177
 QY 61 TCGTTGACAGAAATCCCAATACCGAGCTTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 178 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 197
 QY 120 TTTCAGGGGAGACCCAGTGTCTGGCCAAATTCGACGTCCTCCACCTCCATCA 179
 Db 198 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 217
 QY 180 CTACCAACCTCTGTCTCCAAATTCCTCGGCTATCGTGGATGTCTGGCGGCTTT 239
 Db 217 sSerProThrSerCysProPheCysProGlyTyArgTrpMetCysLeuArgArgPh 237
 QY 240 TATCATATCTCTTCATCTCTGCTATGCTCATCTTCTGTGTGTTCTTCTGGACTA 299
 Db 237 eillePheLeuPheilleLeuLeuLeuCysLeuillePheLeuLeuValLeuLeuAspTy 257
 QY 300 CCAGGTATGTCCTGCTGCTCTTACTTCCAGGAACATCACACACGACGGGGCC 359

Db 257 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 277
 QY 360 ATGCAAGACCTGCACGACTCTGCTCAAGAACTCTAGTTTCCTCTCTGTTGCTGTAC 419
 Db 277 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCys 297
 QY 420 AAAAACTTCGGACGGAAACTGCACCTTGATTCCCATCCCATCATCTCGGCTTTTCGCAAG 479
 Db 297 rIysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 317
 QY 480 ATTCTATATGGAGTGGGCTCAGTCCGTTTCTCTCGGCTCAGTTTACTAGTGCATTTGT 539
 Db 317 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 337
 QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTGGCTTTTTCAGTTATATGATGATGGTA 599
 Db 337 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 357
 QY 600 TTGGGGCGGAGTCTGTACAACTTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
 Db 357 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 377
 QY 660 TTGCTTTTGGGTATACATT 678
 Db 377 eCysLeuTrpValTyIle 383
 RESULT 12
 ADH77843
 ID ADH77843 standard; protein; 406 AA.
 AC ADH77843;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Hepatitis B virus surface antigen, HBsAgL, SEQ ID 2.
 XX
 KW Nanoparticle; surface antigen protein; HBsAgL; liver; drug delivery.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004002459-A1.
 XX
 PD 08-JAN-2004.
 XX
 PF 27-JUN-2003; 2003WO-JP008244.
 XX
 PR 28-JUN-2002; 2002JP-00191386.
 PR 27-JUN-2003; 2003JP-00183863.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Kuroda S, Tanizawa K, Kondo A, Ueda M, Seno M, Tada H;
 XX
 DR WPI; 2004-082876/08.
 DR N-PSDB; ADH77842.
 XX
 PT Hollow nanoparticle comprising protein with modified cysteine residues
 PT for target specific drug delivery.
 XX
 PS Example 1; SEQ ID NO 2; 79pp; Japanese.
 XX
 CC The present invention relates to a hollow nanoparticle, which comprises a
 CC protein, e.g. hepatitis B virus surface antigen protein, with modified
 CC cysteine residues. The nanoparticle specifically recognizes a target cell
 CC e.g. a liver cell. The nanoparticle is useful in methods for drug
 CC delivery. The present sequence is one such hepatitis B virus surface
 CC antigen protein.
 XX
 SQ Sequence 406 AA;
 Alignment Scores:
 Pred. No.: 1.01e-109 Length: 406
 Score: 1206.00 Matches: 220

QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGGAACCTCTAGCTTCCCTCTGTGCTGTAC 419
 Db 300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 320
 QY 420 AAAACCTTCGGACGGAACTGCACCTGTATTCCTCCATCCCATCATCTCGGCTTCGCAAG 479
 Db 320 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 340
 QY 480 ATTCCTATGAGAGTGGGCTCAGTCCGCTTCTCTGCTCAGTTACTAGTGCATTTGT 539
 Db 340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 360
 QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTCGGCTTTCAGTTATATGATGATGGTA 599
 Db 360 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 380
 QY 600 TTGGGGCGGAAGTCTGTACAACTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 380 rTrpGlyProSerLeuTyraAsnIleLeuSerProPheLeuProLeuProIlePhePh 400
 QY 660 TTGCTTTTGGGTATACATT 678
 Db 400 eCysLeuTrpValTyrlle 406

RESULT 14

AD9S1504
 ID ADS91504 standard; protein; 406 AA.

XX AC ADS91504;

DT 16-DEC-2004 (first entry)

DE Amino acid sequence of a HBV surface antigen protein.

XX KW haemophilia; nanoparticle; HBV; surface antigen;

XX KW blood coagulation factor VIII; haemostatic; gene therapy.

XX OS Hepatitis B virus.

XX PN WO2004082720-A1.

XX PD 30-SEP-2004.

XX PF 17-MAR-2004; 2004WO-JP003560.

XX PR 17-MAR-2003; 2003JP-00071788.

XX PA (BEAC-) BEACLE INC.

XX PA (VIBV-) VIB VLAMS INTERUNIVERSITAIR INST BIOTEC.

XX PA (COLL-) COLLEN RES FOUND VZW ONDERWIJZEN NAVORSI.

XX PI Ueda M, Kuroda S, Tanizawa K, Senoo M, Kondo A;

XX PI Vandendriessche T, Chuah M;

XX DR WPI; 2004-699721/68.

XX DR N-PSDB; ADS91503.

XX PT Agent useful for treating hemophilia, comprises gene encapsulated in

XX PT hollow nanoparticle, which is obtained by expressing protein that is

XX PT capable of forming particles in eukaryotic cell.

XX PS Example; SEQ ID NO 2; 33pp; Japanese.

XX The specification describes a haemophilia therapeutic agent, which
 CC comprises a gene encapsulated in a hollow nanoparticle, which is obtained
 CC by expressing a protein that is capable of forming particles in a
 CC eukaryotic cell. The protein that forms the particle is a hepatitis B
 CC virus (HBV) surface antigen protein. The gene is a blood coagulation
 CC factor VIII (IX) gene. Therapeutic agents of the invention are useful for
 CC treating haemophilia. They are efficiently transferred into liver cells
 CC with little risk of side effects. The present sequence represents a HBV
 CC surface antigen protein, which may be used to produce therapeutic agents
 CC of the invention.

XX SQ Sequence 406 AA;

Alignment Scores:

Score: 1.01e-109 Length: 406
 Pred. No.: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.32% Indels: 2
 DB: 8 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x ADS91504 (1-406)

QY 1 ATGGGAAACACACATCAGGATTCCTAGCACCCCTGCTGCTTACAGCGGGGTTTTTC 60
 Db 181 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValGlnAlaGlyPhePhe 200
 QY 61 TCGTTGACAAAGATCCTCAATACCGAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 201 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 220
 QY 120 TTTTAGGGGAGACCCACAGTGTTCCTGGCCAAATTCGCAAGTCCCAACCTCCAAATCA 179
 Db 221 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 240
 QY 180 CTCACCAACCTCTTGTCTCCCAATTTGCTCTGCTATCGCTGATGTGTCTGGCGGTTT 239
 Db 240 sSerProThrSerCysProProIleCysProGlyTyraArgTrpMetCysLeuArgPh 260
 QY 240 TATCATATTCTTTCATCTGCTATGCTATGCTCATCTTCTTGTGTTCTTCTGGACTA 299
 Db 260 llellePheLeuPheilleLeuLeuLeuCysleullePheLeuLeuValLeuLeuAspTy 280
 QY 300 CCAAGGTATGTGCCCGTTTGTCTTACTTCCAGGAACATCAACACCAGCACGGGGCC 359
 Db 280 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 300
 QY 360 ATGCAAGACCTGCAGCAGCTCTGCTCAAGAAACTCTAGTTCCCTCTTGTGCTGTAC 419
 Db 300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 320
 QY 420 AAAACCTTCGGACGGAACTGCACCTGTATTCCTCCATCCCATCATCTCGGCTTTCGCAAG 479
 Db 320 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 340
 QY 480 ATTCCTATGAGAGTGGGCTCAGTCCGCTTCTCTGCTCAGTTACTAGTGCATTTGT 539
 Db 340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 360
 QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTCGGCTTTCAGTTATATGATGATGGTA 599
 Db 360 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 380
 QY 600 TTGGGGCGGAAGTCTGTACAACTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
 Db 380 rTrpGlyProSerLeuTyraAsnIleLeuSerProPheLeuProLeuProIlePhePh 400
 QY 660 TTGCTTTTGGGTATACATT 678
 Db 400 eCysLeuTrpValTyrlle 406

RESULT 15

AEA21247

ID AEA21247 standard; protein; 406 AA.

XX AC AEA21247;

XX DT 11-AUG-2005 (first entry)

XX DE Nanoparticle-associated protein #1.

XX KW nanoparticle; Gene therapy; drug delivery; pharmaceutical.

XX

Unidentified.
WO2005049824-A1.
02-JUN-2005.
19-NOV-2004; 2004WO-JP017282.
21-NOV-2003; 2003JP-00392649.
(NISC-) JAPAN SCI & TECHNOLOGY AGENCY.
Seno M, Tada H, Kuroda S, Tanizawa K, Kondo A, Ueda M;
WPI; 2005-396109/40.
N-PSDB; AEA21246.
Hollow nanoparticle useful in preparation of medical agent for treating
PT disease, made of protein capable of forming particles, and presents
PT biological structure-recognition sites on both N- and C-terminals of
PT protein.
XX Disclosure; Page; 54pp; Japanese.
XX The invention relates to a hollow nanoparticle (I) made of protein
CC capable of forming particles and presents biological structure-
CC recognition sites on its surface, where the biological structure-
CC recognition sites are provided at both the N- and C-terminals of the
CC protein and recognize different chemical structure of biological origin
CC from each other. Also described are the following: production of (I); and
CC a medical agent (II) comprising a cell introduced with a substance and
CC sealed by (I). In (I), the biological structure recognition site is a
CC target structure recognition site which recognizes a cell that serves as
CC a target. The biological structure recognition site functions as an
CC identification region for recognizing (I) specifically. The biological
CC structure recognition site is a receptor of host cell recognition
CC structure derived from virus e.g. hepatocyte recognition site of
CC hepatitis B virus surface antigenic protein, antigenic determinant e.g.
CC tag sequence such as Strep-tagII, His-tag, HA-tag or FLAG-tag, where one
CC or more tag sequences are combined, or ligand e.g. cell growth factor or
CC ZZ-tag. The cell growth factor is epidermal growth factor fibroblast
CC growth factor. The biological structure recognition site combines a tag
CC sequence and ligand. The protein which has particle formation ability is
CC a surface antigen protein derived from virus, preferably hepatitis B
CC virus surface antigen protein. The antigenic determinant and/or ligand
CC are provided as the identification region for recognizing (I),
CC specifically. The biological structure recognition site is introduced
CC into the protein by substituting or adding amino acid sequence of the
CC protein capable of forming particle at any one of the terminal. In (II),
CC the substance introduced into the cell is a compound comprising a gene
CC having pharmacological effect. (I) is useful in the preparation of a
CC medical agent which is useful for treating a disease. (I) is useful as a
CC reagent e.g. nanocapsule in pharmaceutical industry for drug delivery.
CC (I) can be purified efficiently and can be introduced into a cell for
CC preparing a drug. The present sequence represents a nanoparticle-
CC associated protein of the invention. Note: This sequence is not shown in
CC the specification but was obtained in electronic format directly from
CC WIPO at ftp.wipo.int/pub/published_pct_sequences/02.06.2005/.
XX
SQ Sequence 406 AA;

Alignment Scores:
Pred. No.: 1,01e-109 Length: 406
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 9 Gaps: 0
US-10-761-006a-1_copy_155_835 (1-681) x AEA21247 (1-406)
QY 1 ATGGAGAACACAAATCAGGATTCCTAGGACCCCTCTCGTGTATACGGCGGGTTTTC 60
|||||

181 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 200
QY 61 TCGTTGACAGAAATCTCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
|||||
201 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 220
Db
120 TTTCTAGGGGAGCACCACCGTGTTCCTGGCCAAATAATTCGAGTCCCCCAACTCCAATCA 179
|||||
221 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 240
Db
180 CTCACCAACCTCTGTCTCTCAATTTGCTGGCTATCGCTGGATGTCTCTCGCGCGTTT 239
|||||
240 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 260
Db
240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGGTTCCTCTGGACTA 299
QY
260 ellelePheLeuPheleleLeuLeuLeuCysleullePheLeuLeuValLeuLeuAspTy 280
Db
300 CCAAGGTATGTTGCCCGTGTTCCTCTACTTCCAGGAAACATCAACCACGACGCGGGCC 359
|||||
280 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 300
Db
360 ATGCAAGACCTGCGACGACTCTCTGCTCAAGGAAACTCTAGGTTTCCTCTTGTGTCTGTAC 419
QY
300 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 320
Db
420 AAAACCTTCGGACGGGAAACTGCACCTTGATTTCCCATCCATCATCTCTGGCTTCGCAAG 479
QY
320 rLysProSerAspGlyAsnGlyThrCysIleProlleProSerSerTrpAlaPheAlaAr 340
Db
480 ATTCTATGGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCCAATTTGT 539
QY
340 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 360
Db
540 TCAGTGGTTCGTAGGGCTTTCCCCACACTGTTGGCTTTTCAGTTATATGATGATGTGTA 599
QY
360 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 380
Db
600 TTGGGGGGGAGTCTGTACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
QY
380 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProllePhePh 400
Db
660 TTGCTCTTTGGGTATACATT 678
QY
400 eCysLeuTrpValTyIle 406
Db

Search completed: December 27, 2005, 20:51:57
Job time : 187.238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 1, 2005, 23:16:00 ; Search time 94.5 Seconds
(without alignments)
5625.446 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACCAAG.....CCTCAGGCCACGACGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10761006/runat_01122005_113953_26895/app.query.fasta_1.3399
-DB=Issued Patents AA -QFMT=fastan -SURFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODES=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CGN 1 1.140 @runat_01122005_113953_26895 -NCFU=6 -ICPU=3
-NO WMAP -LARGUEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/7 COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/8 COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/9 COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	2	US-09-719-528A-2
2	2927	49.6	843	2	US-10-209-264-2
3	2807	47.5	843	2	US-08-591-502B-45
4	2807	47.5	843	2	US-08-591-502B-59
5	2788	47.2	843	2	US-08-591-502B-47
6	2773	46.9	843	2	US-08-591-502B-46
7	2759	46.7	843	2	US-08-591-502B-48
8	2759	46.7	845	1	US-08-416-950-11
9	2759	46.7	845	1	US-08-469-830-11
10	2741	46.4	845	2	US-08-591-502B-11
11	2734.5	46.3	842	2	US-08-591-502B-51
12	2729.5	46.2	842	2	US-08-591-502B-50

13	2652	44.9	832	2	US-08-591-502B-61
14	2646	44.8	832	2	US-08-591-502B-62
15	2645	44.8	843	2	US-08-591-502B-56
16	2643	44.7	832	2	US-08-591-502B-60
17	2643	44.7	843	2	US-08-591-502B-55
18	2637	44.6	832	2	US-08-591-502B-64
19	2633	44.6	843	2	US-08-591-502B-53
20	2615	44.3	843	2	US-08-591-502B-54
21	2608	44.2	845	2	US-08-591-502B-57
22	2607	44.1	832	2	US-08-591-502B-63
23	2607	44.1	845	2	US-08-591-502B-58
24	2607	44.1	845	6	5196194-17
25	2179	36.9	730	2	US-08-591-502B-49
26	1496	25.3	400	2	US-09-719-528A-3
27	1496	25.3	400	2	US-10-209-264-3
28	1427	24.2	277	6	5164485-2
29	1423	24.1	281	2	US-09-247-890-10
30	1423	24.1	281	2	US-09-724-969-10
31	1423	24.1	281	2	US-09-724-852-10
32	1365.5	23.1	481	2	US-08-591-502B-52
33	1359	23.0	281	1	US-08-105-483-214
34	1359	23.0	281	1	US-08-709-209-214
35	1359	23.0	281	1	US-08-458-101-214
36	1359	23.0	389	1	US-08-105-483-216
37	1359	23.0	389	1	US-08-105-483-219
38	1359	23.0	389	1	US-08-709-209-216
39	1359	23.0	389	1	US-08-709-209-219
40	1359	23.0	389	1	US-08-458-101-216
41	1359	23.0	389	1	US-08-458-101-219
42	1354	22.9	281	2	US-09-247-890-12
43	1354	22.9	281	2	US-09-724-969-12
44	1354	22.9	281	2	US-09-724-852-12
45	1354	22.9	281	2	US-09-721-480-3

ALIGNMENTS

RESULT 1
US-09-719-528A-2
; Sequenced, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USSES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ladass & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/719,528A

APPLICATION NUMBER: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Mass, Clifford J.

REGISTRATION NUMBER: 30,086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-719-528A-2

Alignment Scores:
Pred. No.: 5,098-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-09-719-528A-2 (1-843)

QY 1 CTCACAAACATTCACCAAGCTCTGCTAGATCCACAGGCTGAGGGCCCTATATTTTCCTGC 60
DB LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGAAACAGTAAACCTGTTCGGACTACTGCTCTCCCATATCGTC 120
DB TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGACCCCTGCACCAACATGAGACACCAACATCAGATTCT 180
DB AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTGTACAGGGGGGTTTTCTCGTTGACAAGATCCCTCAACAATCC 240
DB ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTGTGGACTTCTCAATTTCTAGGGGAGCACCCACGCTGCC 300
DB AlaGluSerArgLeuTrpTrpTrpSerLeuAsnPheLeuGlyGlyAlaProThrCysSer 403
QY 301 TGGCCAAATTCGGAGTCCCAACCTCCAACTCACCACCTCTGTCTCTCCAAATTG 360
DB TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLeu 423
QY 361 TCCTGGCTATCGCTGGATGTCTCGCGGCTTTTATCATATTCTCTTCATCCTCGCTGCT 420
DB SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
QY 421 ATGCTCATCTCTTGTGGTCTTCTGGACTACCAAGGTATGTTGGCCGTTTGTCTCT 480
DB MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
QY 481 ACTTCCAGAACATCAACACACGCGGCGCCATGCAAGACCTGCACGACTCTCTGTCA 540
DB ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
QY 541 AGGAACTCTACGTTTCCCTCTTGTGTCTGTACAAAACCTTCGGACGGAAACTGCACCTG 600
DB ArgLysLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
QY 601 TATTCCTCCATCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCCG 660
DB TyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
QY 661 TTTCTCTGGCTCAGTTTACTAGTGGCAATTTTTCAGTGGTTCGTAGGGCTTCCCCCAC 720
DB PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
QY 721 TGTGTGGCTTTTCAGTTATATGATGATGTGTATTTGGGGCGAAAGTCTGTACAACATCTT 780
DB CysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeu 563
QY 781 GAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCCCT 840
```

```
DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
QY 841 AATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGCGATATGTAATGGAGTTGG 900
DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
QY 901 GGTACTTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTTCGAAAACTGCCT 960
DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
QY 961 GTAAATAGACTATGTATGGAAGATGTCAAGAAATGTGGGTCTTTTGGGCTTTGCT 1020
DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
QY 1021 GCCCTTTTACACAAATGTGGCTATCCCTGCTTGTATGCTTTATATGATGATATACATCT 1080
DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
QY 1081 AAGCAGGCTTTCACTTTCTCGCCAACTTACAAGGCCCTTCTGTGTAAACAATATCTCAAC 1140
DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
QY 1141 CTTTACCCCGTGGCCCGCAACCGTCCGCTCTCTGCAAGTGTGTGTGACGCAACCCCT 1200
DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
QY 1201 ACTGGATGGGCTGGCCATAGCCCATCAGGCATGCGCATGCTGGACCTTTCTGGCTCCTCTG 1260
DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
QY 1261 CCGATCCATACTCGGAACCTCTAGCAGCTTGTGTCTCGCAGCCGCTCTGGAGCAAAA 1320
DB 724 ProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1321 CTTATCGGAACCGCAAACTCTGTTGTCTCTCTCGGAATACACCTCTCTTTCAT 1380
DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
QY 1381 CTAGGGTGTCTGCCAATCGATCTGCGGGCGCTTGGGGCTTACCGCTCCCTCTTTCAT 1440
DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
QY 1441 CTGAATCCCGCGACGACCCGCTCTCGGGGCGCTTGGGGCTTACCGCTCCCTCTTTCAT 1500
DB 784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
QY 1501 CTGCGGTTCCGGCGACACCGCGGCGCACTCTCTTTTACCGGGTCTCCCGTATGTGCCT 1560
DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
QY 1561 TCTCATCTCGCGGACCGTGTGCACTTCGCTTCCCTACCTCTGCGATCGCATGAGACCCG 1620
DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaIleTrpArgPro 843

RESULT 2
US-10-209-264-2
; Sequence 2, Application US/10209264
; Patent No. 6781142
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
; USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Ladas & Party
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

```

;
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION DATA:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-209-264-2

Alignment Scores:
Pred. No.: 5,09e-259 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-10-209-264-2 (1-843)
QY 1 CTCACAAATCCACCAAGCTCTGTAGATCCCGAGGTGAGGGCGCTATATTTCTCTGC 60
DB 304 LeuHisAsnIleProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
QY 61 TGTGTGCTCCAGTCCGACAGTAACCTGTTCGACTACTGCTCTCCATATCGTC 120
DB 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGGACCTGCACCGAACATGGAGAACACAAATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTCGTGTTCAGCGGGGTTTTCTCGTTGACAGAATCCTCACAAATACC 240
DB 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGACTAGACTGTGTGGACTTCTCAATTTCTAGGGGGAGCACCGGTGTTCC 300
DB 384 AlaGluSerArgLeuTrpTrpThrSerLeuAsnPheLeuGlyGlyAlaProThrCysSer 403
QY 301 TGGCCAAATTCGAGTCCCACTCCCAATCACTCACCAACTCTGTGCTCCCAATTG 360
DB 404 TrpProIlePheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
QY 361 TCCTGCTATCGCTGGATGTCTGCGGGGTTTTATCATATTCCTTCTCATCTGCTGCT 420
DB 424 SerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAla 443
QY 421 ATGCTCATCTTCTGTGTGTTCTTCTGACTACCAAGTATGTGCTGCTGCTGCT 480
DB 444 MetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSer 463
QY 481 ACTTCCAGAACATCAACACACGAGCAGGGGCCCAAGACCTGCACGACTCCTGCTCA 540
DB 464 ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483

```

```

QY 541 AGGAAACTCTACGTTTCCCTCTTGTCTGTACAAAACCTTCGACGGAACCTGCACCTG 600
DB 484 ArgLysLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
QY 601 TATTTCCCATCCATCATCTCGGCTTTTCGCAAGATTTCTATGGAGTGGGCTCAGTCCG 660
DB 504 TyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
QY 661 TTTCTCTGCTCAGTTTACTAGTGCATTTGTTTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
DB 524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHis 543
QY 721 TGTGTGCTTTTCACTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACACATCTT 780
DB 544 CysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeu 563
QY 781 GAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGGTATATACATTTAAACCT 840
DB 564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
QY 841 AATAAAACCAACGTTGGGCTACTCCCTTAACTTCAATGATGATGATGATGATGATGATG 900
DB 584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
QY 901 GGTACTTTACCGCAGGACATATTGTACTAAACTCAAGCAATGTTTTCGAAACCTGCCT 960
DB 604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
QY 961 GTAAATAGACCTATTGATTGAAAAGTATGTCAAAGAATTTGGGGTCTTTTGGGCTTCT 1020
DB 624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
QY 1021 GCGCTTTTACCAATGTGGCTATCTGCTGCTGATGATGATGATGATGATGATGATGATG 1080
DB 644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
QY 1081 AAGCGGCTTTTCACTTTCTCCCAACTTACAGGCTTCTGTGTAAACAATATCTGAC 1140
DB 664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
QY 1141 CTTTACCCTGTTCCCGCAACGGTCCGCTCTCTGCCAAGTGTTCGTGAGCAACCCCT 1200
DB 684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
QY 1201 ACTGGATGGGCTTGGCCATAGCCATCAGCGCATGCTGGAACCTTTCTGCTCTCTG 1260
DB 704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
QY 1261 CGATCCATCTAGCGGAACCTCTAGCAGCTTGTTCGTCAGCGGCTCTGGAGCAAAA 1320
DB 724 ProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1321 CTTATCGGAACCGCAACTCTGTTGTCTCTCGAAATACACCTCTCTTCATGGCTG 1380
DB 744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeu 763
QY 1381 CTAGGCTGTGCTGCCAATCTGGATCTGCGCGGAGCTCTTTTGTCTACGTCCTCCGTCGG 1440
DB 764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
QY 1441 CTGAATCCCGGAGACGACCGCTCTCGGGCGGCTTTGGGGCTCTACCGTCCCTCTTCTAT 1500
DB 784 LeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
QY 1501 CTGCGGCTTCCGCGGACCGACCGGGCGACCTCTCTTTACGGGCTCTCCCGTATGCT 1560
DB 804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
QY 1561 TCTCATCTGCGGAGCCGCTGTGCATTCGCTTCCCTCTGCGAGTCTGCGATGGAGACCG 1620
DB 824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

```

RESULT 3

US-08-591-502B-45
; Sequence 45, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-08-591-502B-45

Alignment Scores:
Pred. No.: 5e-248 Length: 843
Score: 2807.00 Matches: 528
Percent Similarity: 98.33% Conservative: 3
Best Local Similarity: 97.78% Mismatches: 8
Query Match: 47.52% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-45 (1-843)

QY 4 CACAACTTCCCAAGCTCTGCTAGATCCAGGGTGAGGGCCCTATTATTTCTGCTGG 63
Db 305 HisAenilleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCysTrp 324
QY 64 TGGCTCCAGTTCGGAAAGTAAACCTGTTCCGACTACTGCTCTCCCATATGTCAT 123
Db 325 TrpLeuGlnPheArgAsnSerIysProCysSerAspTyrCysLeuThrHisIleValAsn 344
QY 124 CTTCCTCAGGAGCTGGGGACCTCGACCGAAACATGGAGAACCAACATCAGGATTCCTAGG 183
Db 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364

QY 184 ACCCTCTCTCGTGTTCAGGGCGGGTGTTCCTGCTGACAAAGATCTCTCAATACCGCA 243
Db ThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
QY 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGGAGCACCGTGTTCCTG 302
Db GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
QY 303 GCCAAAATTCGCAGTCCCAACCTCCAATCACCTCACCAACCTCTTGTCTCCCAATTTGTC 362
Db pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
QY 363 CTGGCTATCGCTGAGATGTCCTGCGGCTTTTATCATATTCCTCTCATCTGCTGCTAT 422
Db rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
QY 423 GCCTCATCTTCTTGTGTTCTTCTGAGACTACCAAGTATGTTGCCGTTTGTCTCTCTAC 482
Db tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
QY 483 TTCCAGGAACATCAACCCAGCAGCGGGGCATGCAAGACTGCACGACTCTCTGCTCAAG 542
Db rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
QY 543 GAAACTCTACGTTTCCCTCTTGTGCTGTGTAACAAACCTTCGAGCGGAAATCGACTTGTGA 602
Db gAsnLeuTyr-ValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 504
QY 603 TTCCCATCCCATCATCTCGGGCTTCGCAAGATCTCTATGGAGTGGCGCTCAGTCCGTT 662
Db rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
QY 663 TCTCTCGCTCAGTTTACTAGTGCATTTGTTTCAGTGTGTTCTGAGGCTTTCCCGGCTG 722
Db eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544
QY 723 TTTGGCTTTTCAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
Db sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
QY 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 842
Db uSerLeuPheThrSerIleThrAsnPheLeuLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
QY 843 TAAACCAAAACGTTGGGGCTACTCCCTTAACTTTCATGGGATATGTAATTTGGAAGTTCGGG 902
Db nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 604
QY 903 TACTTTACCGAGAACATATTTGATCTAAAACCTCAAGCAATGTTTTCGAAAACCTGCTGT 962
Db yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
QY 963 AATATAGACCTATTTCATTTGGAAGTATGTCAAAGATTTGCTGGTCTTTTGGCTTTGCTGC 1022
Db lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
QY 1023 CCCTTTTACACAATGTGGCTATCTGCTTGTGAGCTCTTATATGTCATGATATACAATCTAA 1082
Db aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 664
QY 1083 GCAGGCTTTTCATTTCTCGCCAACTTACAAAGGCTTTCTGTGTAAACAATATCTGAACCT 1142
Db sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsnLe 684
QY 1143 TTACCCCGTTCGCGGCAACGCTCGGCTCTCTGCCAGTCTTGTGCTGACGCAACCCCGCAC 1202
Db uTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProTh 704
QY 1203 TGGATGGGGCTTGGCCCATAGCCATCAGCGCATGGCTGGGAACCTTTCTGCGCTCTCTGCC 1262
Db rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
QY 1263 GATCATCTACTCGCGGAACCTCTAGACGCTTGTGTTGCTCGCAGCGCGGTCTGGAGCAAACT 1322

Db 724 oilleHistrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLeu 744
QY 1323 TATCGAACCGCAACTCTGTGTCTCTCTCGAAATACACTCTCTTCCATGCTGCT 1382
Db 744 uileGlyThrAspAsnSerValValSerArgLysThrSerPheProTrpLeu 764
QY 1383 AGGTTGTCTGCAACTGATCCTCGCGGAGACGTCTTTGTCTACGTCCGTCGCGCT 1442
Db 764 uGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTrpValProSerAlaLe 784
QY 1443 GAATCCCGGAGCGCGCTCTCGGGCGGTCTTGGGCTCTACCGTCCCTCTTCACT 1502
Db 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTrpValProLeuLeu 804
QY 1503 GCGTTTCCGCGCGCACCGGCGGACCTCTCTTACGGGTCTCCCGGTATGCTTC 1562
Db 804 uProPheArgProThrThrGlyArgThrSerLeuTrpValValSerProSerValProSe 824
QY 1563 TCATCTGCGGACCGGTGCTCACTTCCTTCACTCTGCACTGCGATGAGACCG 1620
Db 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843

RESULT 4
US-08-591-502B-59
; Sequence 59, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-08-591-502B-59

Alignment Scores: 5e-248 Length: 843
Pred. No.: 2807.00 Matches: 527
Score: 98.34% Conservative: 5
Percent Similarity: 97.41% Mismatches: 8
Best Local Similarity: 47.52% Indels: 2
Query Match: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-59 (1-843)
QY 1 CTCCACAACTTCCACCAAGCTCTGTAGATCCCAAGGTGAGGGGCTATATTTTCCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCCGGAAACAGTAAACCTCTGTTCCGACTACTGCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTGAGGACTGGGGACCTTGCACCGAATGAGGAGACACAAATCAGGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGTTTACAGGGCGGTCTTCTGTTGACAAAGATCCTCACAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCACTCAATCACTCACTCACTCACTCTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTCTGCTATGCTGGATGTCTGCGCGGTCTTATCATATCTCTCTCTCTCTCTCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTrpHisIleProLeuHisProAla 443
QY 420 TATGCTCATCTCTTGTGTTCTCTGAGTACCAAGGTATGTTGCCGTTGCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACACGAGCAGCGGGCCATGACAGCTGACGACTCTCTGCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 483
QY 540 AAGGAACCTCTACGTTTCCCTCTGTTGCTGTACAAACCTTTCGGACGGAATCTGCACIT 599
Db 483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCTATGGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTGCTGCTAGTTTACTAGTGCATTTGTCAGTGGTTCGTAGGGCTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACCGTTGGGCTACTCTTAACTTTCATGGATATGTAAATGGAAGTTG 899
Db 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTTGACTTAAACTCAGCATGTTTTCGAAACTGCC 959

Db 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
Qy 960 TGTAATAGACCTATTGATTGGAAGTATGCAAGAATTTGCGTCTTTTGGGCTTTC 1019
Db 623 oValAsnArgProIleAspTlpYsValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
Qy 1020 TGCCCTTTTACACATGTCCTATCCCTGCTTATGCTTATATGCTATGATATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
Qy 1080 TAAGCAGGCTTTCACTTCTCGCAACTTACAGGCTTCTGTGTAAACAATATCGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheUeuCysGlnTyrLeuAs 683
Qy 1140 CTTTACCCCGTTCGCGCAACGCTCGGTCTCTGCAAGTGTGTTGTCGACGCAACCCC 1199
Db 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
Qy 1200 CACTGATGGGCTTGGCCATAGGCATCAGGCATGCTGGAACCTTCTGCTCTCT 1259
Db 703 oThrGlyTlpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 723
Qy 1260 GCCGATCCATCTCGGAACCTCTAGCAGCTTGTTCGTCGACGCGTCTGGAGCAA 1319
Db 723 uProlleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLy 743
Qy 1320 ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAAATACACCTCTTTCATGCT 1379
Db 743 sLeuileGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTlpLe 763
Qy 1380 GCTAGGCTGCTGCTCAACTGATCTCGCGGAGCGTCTTGTCTACGTCCCGTGGC 1439
Db 763 uLeuGlyCysAlaAlaAsnTlpLeuLeuArgGlyThrSerPheValTyrValProSerAl 783
Qy 1440 GCTGAATCCCGGAGCAGCGCTCTCGGGCCCTTTGGGGCTCTACCGTCCCTTCTCA 1499
Db 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuH 803
Qy 1500 TCTCGCTTCGCGGCGACACGCGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db 803 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
Qy 1560 TTCTCATCTCGCGGACCGTGTGCACTTCGCTTACCTCTGACGTGCGATGAGACACCC 1619
Db 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTlpArgProPr 843
Qy 1620 G 1620
Db 843 o 843

RESULT 5

US-08-591-502B-47
; Sequence 47, Application 10/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B

; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauwer
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-08-591-502B-47

Alignment Scores:

Pred. No.: 2,75e-246 Length: 843
Score: 2788.00 Matches: 523
Percent Similarity: 97.97% Conservative: 7
Best Local Similarity: 96.67% Mismatches: 10
Query Match: 47.20% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-47 (1-843)

Qy 1 CTCACAACTTCACCAAGCTCTGCTAGATCCAGGGTCCAGGGCCCTATATTTCTCTGC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 323
Qy 61 TGTGGTCTCCAGTTCGGAACAGTAAACCTGTTCGACTACTGCTCTCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTCTCCGAGGACTGGGACCCCTCCACCGACATCGAGACACACACATCAGGATTCCT 180
Db 344 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTCGTGTTCACAGCGGGGTTTTCTCGTTGACAGAATCTCTCAATATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGTCTAGACTCTG-GTGGACTCTCTCAATTTTCTAGGGGAGACCCAGTGTTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
Qy 300 CTGCGCAAAATTCGAGTCCCCACCTCCCAATCACTCACCAACCTCTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTCGCTATCGCTGATGTCTGCGGGGTTTTATCATATTTCTCTCATCTCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTCTGTTGTTCTTCTGACTACCAAGGTATGTTCCCGCTTGTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGACATCAACACGACGCGGGCCATCGACAGCTCGACCTCTCTCTCTCTC 539
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483

QY 540 AAGGAACTCTACGTTTCCCTCTGTGCTGTACAAAACCTTCGACGGAACCTGCACTT 599
 Db |||||
 QY 483 rArgAsnLeuYrValSerLeuLeuLeuYrLysThrPheGlyArgLysLeuHisLe 503
 Db |||||
 QY 600 GTATTCCCATCCATCATCTGGGCTTTCGCAAGATTCTATGGAGTGGGCTCAGTCC 659
 Db |||||
 QY 503 uYrSerHisProIleIleuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
 Db |||||
 QY 660 GTTCTCTCGCTCAGTTTACTAGTCCCAATTTGTTGAGTGGTTCGAGGCTTTCCCA 719
 Db |||||
 QY 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValAlaArgAlaPheProHi 543
 Db |||||
 QY 720 CNGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAAGTCTGTACACATCT 779
 Db |||||
 QY 543 sCysLeuAlaPheSerTyMetAspValValLeuGlyAlaLysSerValGlnHisLe 563
 Db |||||
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db |||||
 QY 563 uGlsrLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
 Db |||||
 QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
 Db |||||
 QY 583 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyTrValIleGlySerTr 603
 Db |||||
 QY 900 GGGTACTTTACCGCAGGAAATATGTTACTATAAACTCAAGCAATGTTTCGAAACCTGCC 959
 Db |||||
 QY 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
 Db |||||
 QY 960 TGTAAATAGACTATTGATTGGAAGTATGTCAGAATGTTGGGTCTTTTGGCTTTCG 1019
 Db |||||
 QY 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuGlyPheAl 643
 Db |||||
 QY 1020 TGCCCCCTTTTACACATATGCGCTATCTCGCTTGTATGCTTATATGATGATATCAATC 1079
 Db |||||
 QY 643 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 663
 Db |||||
 QY 1080 TAAGAGGCTTTCACCTTCTCGCAACTTCAAGGCTTCTGTGTAAACAATATCTGAA 1139
 Db |||||
 QY 663 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysGlnGlnTyLeuHi 683
 Db |||||
 QY 1140 CTTTACCCCGTTGCGCGCAAGGTCGCTCTCGCAAGTGTGTTGTCGACGCAACCC 1199
 Db |||||
 QY 683 sLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
 Db |||||
 QY 1200 CACTGATCGGCTTTCGCGCATAGCCATCAGCGCTGCTGGAACCTTCTGCTCTCTCT 1259
 Db |||||
 QY 703 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValValProLe 723
 Db |||||
 QY 1260 GCGATCCATACCTGCGGAACCTCTAGCAGCTTGTGTTGTCGCGCGCTCTGGAGCAAA 1319
 Db |||||
 QY 723 uProIleHisThrAlaGluLeuAlaCysPheAlaArgAspArgSerGlyAlaLys 743
 Db |||||
 QY 1320 ACTTATCGGAACCGACACTCTGCTCTCTCGGAATACACTCTCTTCCATGGCT 1379
 Db |||||
 QY 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 763
 Db |||||
 QY 1380 GCTAGGCTGTGTCGCAACTGGAATCTGCGCGGAGCTCTTGTCTAGCTCCGCTCGGC 1439
 Db |||||
 QY 763 uLeuGlyCysAlaAlaAsnTrpIleuArgGlyThrSerPheValTyValProSerAl 783
 Db |||||
 QY 1440 GCTGAATCCCGCGGACGACCCGCTCTCGGCGCTTGTGGGCTCTACCGCTTCTTCA 1499
 Db |||||
 QY 783 aLeuAsnProAlaAspAspProSerArgLysGlyLeuGlyLeuTyArgProLeuLeuSe 803
 Db |||||
 QY 1500 TCTGCGGTTCCGCGGACGCGGCGACCTCTCTTACGCGCTCTCCCGTATGTCG 1559
 Db |||||
 QY 803 rLeuProPheGlnProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 823
 Db |||||
 QY 1560 TTCTCATCTCGCGGACCGTGTGCACTTCTGCTTCTCACCTCTGACGCTCGCATGGAGCAC 1619
 Db |||||
 QY 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
 Db |||||

QY 1620 G 1620
 Db 843 o 843
 RESULT 6
 ; Sequence 46, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762
 ; REFERENCE/DOCKET NUMBER: 014740-000230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 845 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FEATURE:
 ; NAME/KEY: Modified-site
 ; LOCATION: 16..18
 ; OTHER INFORMATION: /product= "OTHER"
 ; /note= "Xaa = unknown"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
 US-08-591-502B-46
 Alignment Scores:
 Pred. No.: 6,528-245 Length: 845
 Score: 2773.00 Matches: 519
 Percent Similarity: 97.04% Conservative: 6
 Best Local Similarity: 95.93% Mismatches: 15
 Query Match: 46.94% Indels: 2
 DB: 2 Gaps: 0
 US-10-761-006A-1 (1-3215) x US-08-591-502B-46 (1-845)
 QY 1 CTCACCAACATTCACCAAGCTCTGCTAGATCCAGGTGAGGGCTATATTTCTGTC 60
 Db |||||
 QY 306 LeuHisAsnIleProProSerSerAlaArgProGlnSerGlyGlyProIleLeuSerCys 325
 Db |||||

QY 61 TGTGGCTCCAGTTCCGGAAACAGTAAACCCCTGTTCGCACTACTGCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
QY 121 AATCTTCCGAGGACTGGGACCTCGACCCCAACATGAGAGACACACATCAGATTCTT 180
Db 346 AsnLeuLeuGluAAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
QY 181 AGGACCCCTGCTGTGTACAGGGGGGGTTCCTGCTGTGACAAAGATCCCTCACAAATACC 240
Db 366 ArgThrProAlaAsgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTGTTC 299
Db 386 ThrGluSerThrLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 405
QY 300 CTGGCCAAATTCGAGTCCCAACCTCCATCACTACCAACCTCTGTCTCTCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTGGCTATCGTGTGATGTCTGCGGCGTTCATCATATTCCTCTTCATCTCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTCTGTGTGTCTCTCGGACTACCAAGGTATGTGCGCGTTCCTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValValCysLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCAAGACCTGCACGACTCTGCTC 539
Db 465 rThrSerLysAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 485
QY 540 AAGGAAACTTACGTTTCCTCTGTCTGTGTACAAACCTTCGACGCAACCTGCACTT 599
Db 485 rArgAsnLeuTyrValSerLeuPheLeuLeuTyrHisThrPheGlyArgLysLeuHisLe 505
QY 600 GTATTCCCATCCATATCTCTGGGCTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCC 659
Db 505 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTGAGGGCTTTCCCCCA 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 545
QY 720 CTGTTTGGCTTTCAGTTATATGATGTGGTATTTGGGGCGAAGTCTGTACAACTCT 779
Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
QY 780 TGAGTCCCTTTTACCTCTATTACAAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 565 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
QY 840 TAATAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATGTGAATGGAAGTTG 899
Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlyCysTr 605
QY 900 GGGTACTTTACCGAGGACATATTGTACTAAACTCAAGCAATGTTTTCGAAACCTGCC 959
Db 605 pGlyThrLeuProGlnHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 625
QY 960 TGTAAATAGACTATTTGATTCGAAAGTATGCAAGAAATTCGTGGTCTTTTGGGCTTTCG 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
QY 1020 TGCCCTTTTACAAATGTCCTCTGCTTCATGCTTCATGCTTTTATATGCTATGATATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
QY 1080 TAAGCAGGCTTTCACCTTTCTCGCCCAACTTCAAGCGCTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685

QY 1140 CCTTTACCCCGTTCCCGGCAACCGTCCGFTCTCTGCCAGTGTTCGTGACGCAACCCC 1199
Db 685 nLeuTyrProValAlaAargGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATGGGGCTTGCCCATAGCCATCAGCCGATGGCTGGAACTTCTGGCTCCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCCATCCATACTCCGCAACTCTCAGCAGCTTTGTTGCTCGCAGCCGCTCGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 745
QY 1320 ACTTATCGGAACCAACTCTGTGTCTCTCTCGGAAATACACCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
QY 1380 GCTAGGCTGTCTGCCAATCGATCGCTCGCGGGCGCTTTGGGGCTCTACCGTCCCTTCTCA 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
QY 1440 GCTCAATCCCGCGACACCGCTCTCGGGGCGCTTTGGGGCTCTACCGTCCCTTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 805
QY 1500 TCTCGCTTCGCGCGACGACCGCTCTCTTTTACGCGTCTCCCGTATGTGCC 1559
Db 805 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
QY 1560 TTCTCATCTCCGCGACCGTGTGCATCTCGCTTCACCTCTGCACGCTCGACGACGACACC 1619
Db 825 oSerHisLeuProAspArgValHisPheProSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845

RESULT 7
US-08-591-502B-48
; Sequence 48, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762

; REFERENCE/DOCKET NUMBER: 014740-000230US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 48:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 843 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-08-591-502B-48

Alignment Scores:

Pred. No.:	1,25e-243	Length:	843
Score:	2759.00	Matches:	519
Percent Similarity:	97.23%	Conservative:	7
Best Local Similarity:	95.93%	Mismatches:	14
Query Match:	46.71%	Indels:	2
DB:	2	Gaps:	0

US-10-761-006A-1 (1-3215) x US-08-591-502B-48 (1-843)

```
QY 1 CTCACACATTCACCAAGCTCTGTAGATCCCAAGGTGAGGGCTATATTTTCCTGC 60
DB 304 LeuHisAlaIleProProSerCysAlaArgSerGlnSerGluGlyProIleSerSerCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAACCCCTGTTCCGACTACTGCTCTCCATATCGTC 120
DB 324 TrpIleLeuGlnPheArgAsnSerGluProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGACTGGGGACCTCTGACACCAATGAGAACACCAATCAGGATTCCT 180
DB 344 AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCTCTGCTGTTACAGCGGGTTCCTGCTGCTGAGCAAGAACCTCTCAGATACC 240
DB 364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 CGAGAGTCTAGACTGTG-GTGACATCTCTCAATTTCTAGGGGAGCACCCAGCTGTC 299
DB 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY 300 CTGGCCAAATTCGAGTCCCCCACTCCAAATCCTCAATCCTCAACCTCTTGTCTCCAAATT 359
DB 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCTGGCTATCGCTGGATGTCTGCGGGCGTTTATCATATTCCTCTTCATCTCTGCTGC 419
DB 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTTCTTCTGAGTACCAAGTATGTTGCCCGTTTCTGCTCTC 479
DB 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACCAAGCACGGGGCCATGCAAGACCTGCAAGCTCTCTGCTC 539
DB 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACCTCTAGTTTCCTCTTGTGCTGTGTACAAAACCTTCGGACGGAAACCTGACTT 599
DB 483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTTCCCATCCATCCTCGGCTTCGCAAGATTCCTATGAGGTGGGCTCAGTCC 659
DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTTCCTCGGCTCAGTTACTAGTGCATTTGTTTCAGTGGTTCGTAGGGCTTCCCCCA 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 543
QY 720 CTGTTGGCTTTCAGTTATATGAGTATGTTGGGGGCGAAGTCTGTACAAATCT 779
```

```
DB 543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAATCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAAGCTTGGGGTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
DB 583 oHisLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAAACTCAAGCAATGTTTTCGAAAACTGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACCTATTGATTTGAAAGATGATGTAAGAAATTTGGGTCTTTTGGGGCTTGC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACACAATGTCCTATCTGCTTCTGCTTATATGCTTATATGCTATATATATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCATCTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATATCTCAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CCTTACCCCTTCCCGGCAACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1199
DB 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGATGGGGCTTGGCCATAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyGlnSerGlyMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATCTCGGAACTCTTAGCAGCTTGTGCTCGCAGCCGCTCTGAGAGCAAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 743
QY 1320 ACTTATCGGAACCACTCTGTGTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
DB 743 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrple 763
QY 1380 CTAGGGTGTCTGCCAATCTGATCTCGCGGGAGCTCTTGTGCTACGCTCCCTCCGCG 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGGACGACCCGCTCTCGGGGCGCTTGGGGCTCTACCGTCCCTCTTCTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTCGCTTCGCGCGCACCCAGCGGCGCCTCTCTTTTACGGGCTCTCCCGGTATGTGCC 1559
DB 803 sLeuProPheArgProThrThrGlyArgAlaSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCTGCTTACCTCTGCACTGCGATGAGAGCACCC 1619
DB 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
DB 843 o 843
```

RESULT 8

US-08-416-950-11

; Sequence 11, Application US/08416950

; Patent No. 5780036

; GENERAL INFORMATION:

; APPLICANT: CHISARI, Francis V.

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T

; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,950
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE:
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-950-11

```

Alignment Scores:

```

Pred. No.: 1,25e-243 Length: 845
Score: 2759.00 Matches: 516
Percent Similarity: 97.23% Conservative: 10
Best Local Similarity: 95.38% Mismatches: 14
Query Match: 46.71% Indels: 2
DB: 1 Gaps: 0

```

US-10-761-006A-1 (1-3215) x US-08-416-950-11 (1-845)

```

Qy 1 CTCACAACTTCACCAAGCTCTGTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys 325
Qy 61 TGTGGCTCCAGTTCGGAGACAGTAACCCCTGTTCCGACTACTGCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 345
Qy 121 AATCTTCTCGAGGACTGGGACCTCGACCGAACATCGGAGACACACATCAGATTCTCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 365
Qy 181 AGGACCCCTGCTCGTGTGTACAGGGGGGTTTCTCGTTGACAGAACTCTCAAAATACC 240
Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
Qy 241 GCAGAGTCTAGACTCTG-GTGAGCTTCTCAATTTTCTAGGGGGAGACCCACGTTTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
Qy 300 CTGGCCAAATTCGACGTCCCACTCCATCACTCACCAACCTCTTGTCTCCCTCAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425

```

```

Qy 360 GTCCTGGCTATCGCTGGATGTGTCTGGCGGTTTATCATATTTCTCTTTCATCTCTGCTGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 445
Qy 420 TATGCGCTCATCTTCTTGTGGTCTTCTTGGACTACCAAGGTATGTGGCCGCTTGTGCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
Qy 480 TACTTCCAGGAACATCAACACACAGCAGCGGGCCATGCAAGACTGCACGACTCTCTCTC 539
Db 465 rAsnSerArgIleIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
Qy 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 485 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
Qy 600 GTATTGCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db 505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
Qy 660 GTTCTCTCGGCTCAGTTTACTAGTGCATTTGTTCAGTGTGTGTGTGTGTGTGTGTGTGT 719
Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProH 545
Qy 720 CTGTTTGGCTTTCAGTTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
Db 545 sCysLeuAlaPheSerTyrMetAspAspValValLeuLeuGlyAlaLysSerValGlnHisLe 565
Qy 780 TGAGTCCCTTTTACCTCTATATACCAATTTTCTTGTCTTGTGGGTATATATTAACCC 839
Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuLeuSerLeuGlyIleHisLeuAsnPr 585
Qy 840 TAATAAAACCAACCGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAAATTTGGAAGT 899
Db 585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
Qy 900 GGGTACTTTTACCGCAGAACATATTTGTACTTAAACTCAAGCAATGTGTTCGAAACCTGCC 959
Db 605 pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
Qy 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCTCAAGAAATTTGGGCTTTTGGCTTTC 1019
Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
Qy 1020 TGCCCTTTTACACAATGTGCTATCTCTGCTTGTGCTTGTATATCATGTATATACAAATC 1079
Db 645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
Qy 1080 TAAGCAGGCTTTCACCTTCTCGCCAACTTACAAGGCTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
Qy 1140 CTTTACCCTGTTCCCGGCAACGGTCCGCTCTCTGCAAGTGTTCCTCAGCAACCC 1199
Db 685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
Qy 1200 CACTGGATGGGCTTGGCCATAGCCATAGCGCATGGCTGGAACCTTCTGGCTCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
Qy 1260 GCCGATCCATPACTCGGAACTCTTAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1319
Db 725 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 745
Qy 1320 ACTTATCGGAACCGACACTCTGTGTCTCTCTCTCGGAAATATACACTCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
Qy 1380 GCTAGGCTGTGCTCCCAACTGGATCTCTGCGGGGACGCTCTTGTGTCTACGTCCTGCGGC 1439
Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
Qy 1440 GCTGAATCCCGCGGAGACCCGCTCTCGGGGCGGCTTGTGGGCTCTTACCGTCCCTTCTTCA 1499

```



```
QY 1080 TAAGCAGCGTTTCACTTTCTCGCCCACTTACAGCGCTTTCTGTGTAAACAATATCTGAA 1139
Db 665 rLysGlnAlaPheThrPheSerProThrTy rLysAlaPheLeuCy sGlnTy rLeuAs 685
QY 1140 CTTTACCGCTTGC CGGCAACGTCGGCTCTCTGCCAAGTGTTCGTACGCAACCC 1199
Db 685 nLeuTy rProValAlaArgGlnA rProGlyLeuCy sGlnValPheAlaSerAlaThrPr 705
QY 1200 CACTGGATGGGGCTTGCCATAGGCATCAGCGATGGCTGGAACTTCTTGGCTCTCTCT 1259
Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCATCCACTCGGGAACCTCTAGCAGCTTGTGCTCGCAGCGCGTCTGGAGCAAA 1319
Db 725 uProIleHisThrAlaGluLeuLeuAlaAaCy sPheAlaArgSerArgSerGlyAlaLy 745
QY 1320 ACTTATCGAACCCGACAACTCTGTGTCCTCTCTCGGAATACACCTCTTCCATGGCT 1379
Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTy rThrSerPheProTrpLe 765
QY 1380 GCTAGGGTGTCTGCCAACTGGATCCTCGCGGGACGTCCTTTGTCTACGTCCTCGTGGC 1439
Db 765 uLeuGlyCy sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTy rValProSerAl 785
QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTCTTCA 1499
Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgGlyLeuTy rArgProLeuLeuAr 805
QY 1500 TCTCGCTTCCGGCGGACCGACCGGGCGCACCTCTCTTACGCGGTCTCCCGTATGTC 1559
Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTy rAlaValSerProSerValPr 825
QY 1560 TTTCATCTCGCGGACCGTGTGCACCTTCGCTTACCTCTGCACGCTCGCATGGAGACACC 1619
Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db 845 o 845

RESULT 10
US-08-591-502B-11
; Sequence 11, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
```

```
;
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..845
; OTHER INFORMATION: /product= "OTHER"
; /note= "xaa = any amino acid
; (<50% consensus)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-591-502B-11
```

```
Alignment Scores: 5.56e-242 Length: 845
Pred. No.: 2741.00 Matches: 514
Score: 96.30% Conservative: 7
Percent Similarity: 95.01% Mismatches: 19
Best Local Similarity: 46.40% Indels: 2
Query Match: 2 Gaps: 0
DB:
```

US-10-761-006A-1 (1-3215) x US-08-591-502B-11 (1-845)

```
QY 1 CTCACAACTTCACCAAGCTCTGTAGATCCAGGTCAGGGGCTATATTTTCTGTC 60
Db 306 LeuHisAsn***ProProAsnSerAlaArgSerGlnSerGlyProValPheSerCys 325
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCTCTGTCGAGTACTGCTCTCCCATATCGTC 120
Db 326 TrpTrpLeuGlnPheArgAsnSerLysProCy sSerAspTy rCysLeu***HisIleVal 345
QY 121 AATCTTCTCGAGGACTGGGACCTCGACCCGAAATGGAGAACACAAATCAGGATTCCT 180
Db 346 AsnLeuLeuGluAspTrpGlyProCy sThrGluHisGlyGluHis***IleArgIlePro 365
QY 181 AGACCCCTGCTCGTGTACAGCGGGGTTTCTCGTTGACAGATCTCTCAATATACC 240
Db 366 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGACCCACGTCGTC 299
Db 386 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGly***ThrArgVal-Se 405
QY 300 CTGGCCAAAATTCGAGTCCCCAACCTCCAATCACTCACCAACCTCTGTCTCTCCAAATTT 359
Db 405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTGGCTATCGCTGATGTGTCTGGGGGTTTATCATATTCCTCTTCTCATCTGTCGC 419
Db 425 uSerTrpLeuSerLeuAspValSerAlaAlaPheTy rHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCTCATCTTCTTGTGTTCTTCTTGACTACCAAGGTATGTTCGCGGTTTGTCTC 479
Db 445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTy rValAlaArgLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACCAGCAGCGGGGCGCATGCAAGACCTGCACGACTCTGCTC 539
Db 465 rAsnSerArgIleIleAsn***GlnHisGlyThrMetGlnAsnLeuHisAspSerCy sSe 485
QY 540 AAGGAACTCTACGTTTCCCTCTTGTGCTGTACAAACCTTCGGACGGAACTGCACTT 599
Db 485 rArgAsnLeuTy rValSerLeuLeuLeuTy rLysThrPheGlyArgLysLeuHisLe 505
```

QY 600 GTATTCCCATCCATCATCTGGGCTTTCCAGAGATTCCTATGGGAGTGGGCTCAGTCC 659
 Db 505 uTySerHisProIleLeuGlyPheArglysiIleProMetGlyValGlyLeuSerPr 525
 QY 660 GTTCTCTCCGGCTCAGTCTAGTCCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCA 719
 Db 525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 545
 QY 720 CTGTTTGGCTTTTCCAGTTATATGATGATGGTATTTGGGGCGAAGTCTGTACAACATCT 779
 Db 545 sCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
 QY 780 TGAGTCCCTTTTACCTCTATACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db 565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
 QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTATGGGATATGTAATGGAGTTG 899
 Db 585 oAsnLysThrLysArgTrpGlyTySerLeuAsnPheMetGlyTyValIleGlySerTr 605
 QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTTAAACTCAAGCAATGTTTTCGAAACTGCC 959
 Db 605 pGlyThrLeuProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
 QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAGAAATGTCGGTCTTTGGGCTTTGC 1019
 Db 625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuGlyPheAl 645
 QY 1020 TGCCCTTTTACACAATGGCTATCTGCTGCTGATGCTTTATATGATGATGATATCAATC 1079
 Db 645 aAlaProPheThrGlnCysGlyTyProAlaLeuMetProLeuTyAlaCysIleGlnSe 665
 QY 1080 TAAGCAGGCTTTTCACTTTCTCGCAACTTCAAGGCTTTTCTGTGTAACAATATCTGAA 1139
 Db 665 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLeuAs 685
 QY 1140 CTTTACCCCGTTGCGCGCAACGGTCCGGTCTCTGCGAAGTGTGTCGACGCAACCC 1199
 Db 685 nLeuTyProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
 QY 1200 CACTGATGGGCTTCGCCATAGCCATCAGCCATGCGTGGCACTTTCTGCTCCTCT 1259
 Db 705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
 QY 1260 GCGCATCCATCTCGGAACTCTAGCAGCTTGTGTCGCGAGCGGTCTGGAGCAA 1319
 Db 725 uProIleHisThrAlaGluLeuLeuAlaLysPheAlaArgSerArgSerGlyAlaLys 745
 QY 1320 ACTTATCGGAACCGAACAATCTGTTGTCCTCTCGGAATATACCTCTCTTCATGGCT 1379
 Db 745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 765
 QY 1380 GCTAGGCTGCTGCCAAGTGGATCTGCGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
 Db 765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 785
 QY 1440 GCTGAATCCCGGACGACCGCTCTCGGGCGCGTGTGGGCTCTACCGTCCCTCTTCA 1499
 Db 785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyArgProLeuLeuAr 805
 QY 1500 TCTGCCGTTCCGCGCAACCGGGCGCACTCTCTTTTACGCGGTCTCCCGATGTGCC 1559
 Db 805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaValSerProSerValPr 825
 QY 1560 TTCTCATCTCCGACCGTGTGCACTTGGCTTCACTCTGACCTGCGATGCGATGGAGCACCC 1619
 Db 825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
 QY 1620 G 1620
 Db 845 o 845

RESULT 11
 US-08-591-502B-51
 ; Sequence 51, Application US/08591502B
 ; Patent No. 6607727
 ; GENERAL INFORMATION:
 ; APPLICANT: Chisari, Francis V.
 ; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
 ; Lymphocyte Responses to Hepatitis B Virus
 ;
 ; NUMBER OF SEQUENCES: 99
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/591,502B
 ; FILING DATE: 20-May-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,540
 ; FILING DATE: 26-AUG-1991
 ; APPLICATION NUMBER: US 07/935,898
 ; FILING DATE: 26-AUG-1992
 ; APPLICATION NUMBER: US 08/100,870
 ; FILING DATE: 02-AUG-1993
 ; APPLICATION NUMBER: WO PCT/US94/08685
 ; FILING DATE: 01-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Weber, Ellen Lauver
 ; REGISTRATION NUMBER: 32,762
 ; REFERENCE/DOCKET NUMBER: 014740-000230US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 51:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 842 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
 ;
 ; US-08-591-502B-51
 ;
 Alignment Scores:
 Pred. No.: 2,19e-241 Length: 842
 Score: 2734.50 Matches: 517
 Percent Similarity: 96.86% Conservative: 7
 Best Local Similarity: 95.56% Mismatches: 15
 Query Match: 46.29% Indels: 3
 DB: 2 Gaps: 1
 ;
 US-10-761-006A-1 (1-3215) x US-08-591-502B-51 (1-842)
 QY 1 CTCACCAACATCCACCAAGCTCTGTAGATCCAGGCTGAGGGCTATATTTCTCTGC 60
 Db 304 LeuHisHisIleSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
 QY 61 TGTGTGCTCCAGTTCGCGACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrcCysLeuThrHisIleVal 343
 QY 121 AATCTTCTCGAGGACTGGGGACCCCTGCACCGCAACATGGAGAACACCAACATCAGATTCT 180
 Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363

QY 181 AGGACCCCTGCTCGTGTACAGCGGGGGTCTTCTCGTTGACAAAGATCCTCAATACC 240
Db |||||
364 ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG-GTGACATCTCTCAATTTCTAGGGGAGCACCCAGTGTTC 299
Db |||||
384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgLysThrHisVal- 403
QY 300 CTGGCCAAATTCGAGTCCCACTCCCAATCACTCAACCACTCTTGTCTCCTCAATTT 359
Db |||||
403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY 360 GTCCCTGGCTATCGCTGGATGTCTGCGCGTTTATCATATCTCTTCATCTGCTGC 419
Db |||||
423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TATGCTCATCTCTTGTGTGTTCTCTGGACTACCAAGGTATGTGCGCGTTTGTCTC 479
Db |||||
443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaAlaArgLysSer 463
QY 480 TACTTCCAGGAACATCAACACCAACAGCAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db |||||
463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAACTCTACGTTTCCCTCTTGTCTGTGTAACAAACCTTCGACGGAACCTGCACTT 599
Db |||||
483 rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY 600 GTATTCCCATCCATCATCTGGGCTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCC 659
Db |||||
503 uTyrSerHisProIleIleLeuGlyPheArgLysIlePrometGlyGlyGlyLeuSerPr 523
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTCACTGGTTCCTAGGGCTTTCCCCCA 719
Db |||||
523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTGGCTTTCAGTTATATGATGATGTGTATGGGGGGAAGTCTGTACAACTCT 779
Db |||||
543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TGAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db |||||
563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db |||||
583 oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGGAACATATGTACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
Db |||||
603 pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAATAGACTATTGTAAGAGTATGCAAGAAATCTGGGTCTTTTGGGCTTTC 1019
Db |||||
623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACAAATGTGGCTATCTCGCTTCATGCTTTTATATGATCATGTATACATC 1079
Db |||||
643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCATCTTCGCGCACTTACAGGCCCTTCTGTGTAAACAAATATCTGAA 1139
Db |||||
663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi 683
QY 1140 CTTTTPACCCGTTGCGCGCAACGTCGGTCTCTGCCAAGTGTGTGTCGACGCAACCCC 1199
Db |||||
683 sleuTyrProValAlaArg---ArgThrAlaLeuLeuCysGlnValPheAlaAspAlaThrPr 702
QY 1200 CACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGTCTCTCT 1259
Db |||||
702 ofrGlyTrpGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 722
QY 1260 GCCGATCCATACCTGCGGAACCTCTAGCAGCTTGTCTGCGACCGGCTCGGAGCAA 1319

Db |||||
722 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 742
QY 1320 ACTTATCGGAACCGACAACTCTGTGTCTCTCGGAAATACACCTCTCTTTCATGCT 1379
Db |||||
742 sleuIleGlyThrAspAsnSerValLeuSerArgLysTyrThrSerPheProTrpLe 762
QY 1380 GCTAGGTTGTCTGCCAACTGGATCTCGCGGGACGCTCTTGTCTACGTCCTCGTCCG 1439
Db |||||
762 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl 782
QY 1440 GCTGAATCCCGCAGACGACCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTCTTCA 1499
Db |||||
782 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi 802
QY 1500 TCTCGCTTCGGCGCAGCACGCGGCGCTCTTTTACGCGGTCTCCCGTATGTGCC 1559
Db |||||
802 sleuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 822
QY 1560 TTCTCATCTCGCGACCGCTGTGCATCTCGCTTACCTCTGCACCTCGCATCGAGACACC 1619
Db |||||
822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 842
QY 1620 G 1620
Db 842 o 842
RESULT 12
US-08-591-502B-50
; Sequence 50, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 842 amino acids
; TYPE: amino acid
; STRANDBDNSS: <Unknown>

; TOPOLOGY: linear	
; MOLECULE TYPE: protein	
; SEQUENCE DESCRIPTION: SBQ ID NO: 50:	
US-08-591-502B-50	
Alignment Scores:	
Pred. No.:	6,28e-241
Score:	2729.50
Percent Similarity:	96.67%
Best Local Similarity:	95.38%
Query Match:	46.21%
DB:	2
US-10-761-006A-1 (1-3215) x US-08-591-502B-50 (1-842)	
QY	1 CTCACAAACATCCACCAAGCTCTGTAGATCCAGGCTGAGGGCGCTATATTTCTCTGC 60
DB	304 LeuHisHisSerProSerProAlaArgSerGlnSerGluGlyProIlePheSerSer 323
QY	61 TGTGGCTCCAGTCCGGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATGCTC 120
DB	324 TrpTrpLeuGlnPheArgAsnSerLysProCysCysAspTyrCysLeuThrHisIleVal 343
QY	121 AATCTCTCGAGACTGGGACCCCTGCACGACATGGAGAACACACATCAGGATTCCT 180
DB	344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY	181 AGGACCCCTCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCCTCACATACC 240
DB	364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY	241 GCAGAGCTAGACTCTG-GTGGACTCTCTCAATTTCTAGGGGGAGCCACACGTGTTTC 299
DB	384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 403
QY	300 CTGGCCAAATTCGAGTCCCGAACCTCCAACTCACTCACCAACTCTGTCTCCCAATT 359
DB	403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 423
QY	360 GTCTGGCTATCTGCTGATGTGTCTGGCGGTTTTATCATATTTCTCTCATCTCTCTGC 419
DB	423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAla 443
QY	420 TATGCTCTCAPCTTCTTGTGGTCTTCTGCACTACCAAGTATGTTCGCGTGTGCTTC 479
DB	443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY	480 TACTTCCAGGAACATCAACCAACGAGCGGGCCATGCAAGACTGCAAGCTCTCTGCTC 539
DB	463 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY	540 AAGGMAACTCTACCTTTCCCTCTGTGTGTGTACAAACCTTCGGACGGAAACTGCACATT 599
DB	483 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 503
QY	600 GTATTCCTCCATCATCTCTGGCTTTCGCAAGATTCCTATGGAGTGGCGCTCAGTCC 659
DB	503 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr 523
QY	660 GTTCTCTCGCTCAGTTACTAGTGCATTTGTTGAGTGTGTGAGGGCTTCCGCCA 719
DB	523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaAlaPheProHi 543
QY	720 CTGTTTGGCTTTCAGTTATGATGATGTGTATTGGGGCGCAAGTCTGTACAACTCT 779
DB	543 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 563
QY	780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB	563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY	840 TAATAAAACCAAGCTTGGGGCTACTCCCTTAACCTCATCGGATATGTAATTTGGAAGTTG 899

Db	583	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	603
QY	900	GGGTACTTTTACCGCAGGAACATATTTGTACTTAAACTCAAGCAATGTTTTTCGAAACTGCC	959
Db	603	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	623
QY	960	TGTAATAAGACCTATTGATTGGAAAGTATGCAAAAGAAATTTGGGCTCTTTTGGGCTTTGC	1019
Db	623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
QY	1020	TGCCCCCTTTTACCAATGTGGCTATCTTCGCCCTTGATGCTTTTATATGATGATATCAATC	1079
Db	643	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	663
QY	1080	TAAGCAGGCTTTCACCTTCTGCCCAACTTACAGGCCCTTCTGTGTAAACATATCTGAA	1139
Db	663	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi	683
QY	1140	CCTTTTACCCCGTTGCCGCAACGCTCCGCTCTCTGCCAAGTGTGTGTGCTGACGCAACCC	1199
Db	683	uLeuTyrProValAlaArg---ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr	702
QY	1200	CACCTGATGGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGCTCTCT	1259
Db	702	oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe	722
QY	1260	GCCGATCCATATCTGCGGAACCTCTTAGCAGCTTGTGTGTGCTCCAGCGGCTCTGGAGCAA	1319
Db	722	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	742
QY	1320	ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAAATACACTCTTCTCCATGGCT	1379
Db	742	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpue	762
QY	1380	GCTAGGCTGTCTGCGCAACTGATCTCTGCGGGGACGCTCTTGTGTCTAGCTCCGCTCGC	1439
Db	762	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl	782
QY	1440	GCTGAATCCCGCGGACGACCCCTCTCTCGGGGCGCTTGTGGGCTCTACCGTCCCTTCTCA	1499
Db	782	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuIleArgProLeuLeuHi	802
QY	1500	TCTGCGTTCGCGCGGACCGGCGGCGACCTCTCTTACGCGGCTCTCCCGCTATCTGCC	1559
Db	802	sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	822
QY	1560	TTCTCATCTGCGGACCGGTGTCACCTTTCACCTCTGACGCTGCGCATGGAGACCAACC	1619
Db	822	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	842
QY	1620	G 1620	
Db	842	O 842	

RESULT 13

US-08-591-502B-61

Sequence 61, Application US/08591502B

Patent No. 6607727

GENERAL INFORMATION:

APPLICANT: Chisari, Francis V.

TITLE OF INVENTION: Peptides for Inducing Cytotoxic T Lymphocyte Responses to Hepatitis B Virus

NUMBER OF SEQUENCES: 99

CORRESPONDENCE ADDRESS:

ADDRESSES: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,502B
FILING DATE: 20-May-1996
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,540
FILING DATE: 26-AUG-1991
APPLICATION NUMBER: US 07/935,898
FILING DATE: 26-AUG-1992
APPLICATION NUMBER: US 08/100,870
FILING DATE: 02-AUG-1993
APPLICATION NUMBER: WO PCT/US94/08685
FILING DATE: 01-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 014740-000230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 61:

US-08-591-502B-61

Alignment Scores:

Pred. No.: 7,846-234 Length: 832
Score: 2652.00 Matches: 494
Percent Similarity: 95.38% Conservative: 22
Best Local Similarity: 91.31% Mismatches: 24
Query Match: 44.90% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x US-08-591-502B-61 (1-832)

```
QY 1 CTCCACACATTCACCAAGCTCTGTAGATCCAGGCTGAGGCGCTATATTTCCTGC 60
DB 293 LeuHisAsnLeuProAsnSerAlaArgSerGlnSerGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTCCCGAACAGTAAACCCCTGTTCCGACTCTGCTCCCATATCGTC 120
DB 313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisIleVal 332
QY 121 AATCTTCTCGAGGATGGGGACCCCTGCACCGAACATGGAGAACACACATCAGGATTCT 180
DB 333 AsnLeuLeuGluAspTrpGlyProCysAlaGluHisGlyGluHisIleArgIlePro 352
QY 191 AGGACCCCTGCTCGTGTACAGCGGGGTTTTCGTGTGACAGAAATCTTCACAAATACC 240
DB 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 GCAGAGTCTAGACTCTG-GTGACTTCTCAATTTCTAGGGGGAGCACCCAGTGTC 299
DB 373 AlaGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlyAsnTyrArgVal-Se 392
QY 300 CTGGCCAAAATTCGAGTCCCAACCTCCAATCACTCAACCACTCTTTGTCTCTCAATTT 359
DB 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 412
QY 360 GTCCTGGCTATGCTGGATGTCTGCGCGCTTTTATCATATTCTCTTCATCTGCTGTC 419
DB 412 userTrpLeuSerLeuAspValSerAlaAlaPheTyrHisLeuProLeuHisProAlaAl 432
QY 420 TATGCTCATCTCTTGTGTGTTCTTCTGGACTACCAAGGTATGTTGCCGCTGTGCTC 479
DB 432 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 452
```

```
QY 480 TACTTCCAGGAACATCAACCAACGAGCGGGCCATGCAAGACCTGCACGACTCTCTGCTC 539
DB 452 rAsnSerArgIlePheAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 472
QY 540 AAGGAACTCTACGTTTCCCTCTTGTGCTGTACAAAACCTTCGAGCGGAACATCGACTT 599
DB 472 rArgAsnLeuTyrValSerLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCCATCCCATCTCCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
DB 492 uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCTCTGCTCAGTTTACTAGTGCATTTGTTTCACTGTGTTGCTAGGCTTTCCCA 719
DB 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGGTATTTGGGGCGGAGTCTGTACAACTCT 779
DB 532 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
DB 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTTGGGCTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTG 899
DB 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlyCysTy 592
QY 900 GGTACTTTACCGGAGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
DB 592 rGlySerLeuProGlnAspHisIleLeuGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAATAGACCTATTGATTCGAAAGTATGCAAGAAATCTGTGGTCTTTTGGGCTTTC 1019
DB 612 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCTTTTACAAATGTGGCTATCTCTGCTTATGATGCTTTTATATGATGATACATC 1079
DB 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAACGAGCTTTCACTTTCTGCCAACATTACAGGCTTTTCTGTGTAAACAATATCTGAA 1139
DB 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTACCCCTTCCCGCAACGCTCGGCTCTGCCAAGTGTTCGTCAGCGCAACCC 1199
DB 672 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY 1200 CACTGGATGGGCTTGGCCATAGCCCATGCGCATGGCTGGAACCTTTCTGGGCTCTCT 1259
DB 692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaArgLe 712
QY 1260 GCCGATCCATCTCGGGAACCTCTAGCAGCTGTTGTTGCTCGCAGCGGCTCGAGCAAA 1319
DB 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY 1320 ACTTATCGGAACCGACAACTCTGTGTCTCTCGGAATAACACCTCTTCCATGGCT 1379
DB 732 nIleLeuGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerTyrProTrpLe 752
QY 1380 GCTAGGTTGTCTCCCAACTGGATCCTCGCGGGAGCGTCTTGTCTACGCTCCCGTCGGC 1439
DB 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY 1440 GCTGAATCCCGGAGACGACCGCTCTCGGGCGGCTTGGGGCTCTACCGTCCCTTCTTCA 1499
DB 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
QY 1500 TCTCCGCTTCGGGCGGACCGGCGGCTCTCTTTTACGGGCTCTCCCGTATGTGCC 1559
DB 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr 812
```



```
Db 672 nLeuTy-ProValAlaGlyGlnArgProGlyLeuCyseGlnValPheAlaAspAlaThrPr 692
Qy 1200 CACTGGATGGGCTTGCCATAGCCATCAGCGATCGCTGGACCTTTCTGCTCTCTCT 1259
Db 692 oThrGlyTrpGlyLeuAlaMetGlyHisGlnArgMetArgGlyThrPheSerAlaProLe 712
Qy 1260 GCCGATCCACTACTCGGAACTCTCAGCAGCTGTGTTTGTCTCGCAGCGCGTCTGAGCAAA 1319
Db 712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
Qy 1320 ACTTATCGGACCGACCACTCTGTGTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db 732 nIleIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 752
Qy 1380 GCTAGGCTGTGCTGCCCACTGGATCTCTCGCGGACGCTCTTGTCTACCTCCCGTGGC 1439
Db 752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyValProSerAl 772
Qy 1440 GCTGAATCCCGCGGACGACCGCTCTCGGGGCGCTTTCGGGCTCTACCGTCCCTTCTTCA 1499
Db 772 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuSerArgProLeuLeuAr 792
Qy 1500 TCTGCCGCTTCCGCGGACGACCGGCGGACCTCTCTTTTACGCGGTCTCCCGTATGCGC 1559
Db 792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
Qy 1560 TTCTCATCTCGGACCGTGTGCACTCTCGCTTCACTCTCGACGTCGACGTCGACGAGACAC 1619
Db 812 oSerHisLeuProAspLeuValHisPheAlaSerProLeuHisValAlaTrpArgProPr 832
Qy 1620 G 1620
Db 832 O 832
```

RESULT 15

```
US-08-591-502B-56
; Sequence 56, Application US/08591502B
; Patent No. 6607727
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; TITLE OF INVENTION: Peptides for Inducing Cytotoxic T
; Lymphocyte Responses to Hepatitis B Virus
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,502B
; FILING DATE: 20-May-1996
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: WO PCT/US94/08685
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Leuwer
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 014740-000230US
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; MOLECULE TYPE: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-08-591-502B-56
Alignment Scores:
Pred. No.: 3 46e-233 Length: 843
Score: 2645.00 Matches: 493
Percent Similarity: 95.19% Conservative: 22
Best Local Similarity: 91.13% Mismatches: 25
Query Match: 44.78% Indels: 2
DB: 2 Gaps: 0
US-10-761-006A-1 (1-3215) x US-08-591-502B-56 (1-843)
Qy 1 CTCACAAACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTCTGC 60
Db 304 LeuHisValProAsnSerArgSerGlnSerGlnGlySerValLeuSerCys 323
Qy 61 TGGTGGCTCCAGTCCGGACAGTAAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerGluHisCysLeuPheHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCCCTGCACCGAAACATGGAGAACACAAACATCAGATTCCCT 180
Db 344 AsnLeuIleGluAspTrpGlyProCysAlaGluHisGlyGluHisArgIleArgThrPro 363
Qy 181 AGGACCCCTCTCGTGTACAGCGGGGTTTTCTCGTGACAAAGATCTCTCAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGACTAGACTCTG-GTGACTTCTCTCAATTTCTAGGGGGAGCACCACGCTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrArgVal-Se 403
Qy 300 CTGGCCAAAATTGCGAGTCCCCAACCTCCAATCACTCACCAAGCTCTTGTCTCTCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAspLe 423
Qy 360 GTCTGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATTCCTCTTCATCTGCTGC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisLeuProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTCTTGTGTCTGCTGACTACCAAGGTATGTCGCCGTTTGTCTCTC 479
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 463
Qy 480 TACTTCCAGGAACATCAACCAACAGCAGCGGGGCGCATCAAGACCTGCACGACTCTCTGCTC 539
Db 463 rAsnSerArgIleIleAsnHisGlnHisArgThrMetGlnAsnLeuHisAspSerCysSe 483
Qy 540 AAGGAAACTCTTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 483 rArgAsnLeuTyTrValSerLeuMetLeuLeuTyThrLysThrGlyArgLysLeuHisLe 503
Qy 600 GTATTTCCCATCCCATCATCTCTGGCTTTCAGAGATTCCTATGGAGTGGGCTCACTCC 659
Db 503 uTyTrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
Qy 660 GTTCTCTCTGCTCAGTTTACTAGTGCATTGTTCAGTGGTTCGTAGGGGCTTTCCCCCA 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
Qy 720 CTGTTTGGCTTTTCAGTTATATGATGATGTGGTATTTCGGGGCGGAGTCTGTGTACAACTCT 779
```

Db	543	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	563
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTGGGTATACATTTAAACCC	839
Db	563	uGluSerLeuTyrAlaAlaValThrAsnPheLeuLeuSerLeuGlyLeHisLeuAsnPr	583
QY	840	TAATATAAACCAAGCTTGGGGCTACTCCCTTAATCTTCATGGGATATGTAATTCGAAGTTG	899
Db	583	oGlnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	603
QY	900	GGGTACTTACCGCAGCAACATATGTACTATAAATCTCAAGCAATGTTTTCGAAACTGCC	959
Db	603	pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr	623
QY	960	TGTAATAATAGACTTATGATTGGAAAGTATGTCAAAGAAATTTGGGTCTTTTGGGCTTTGC	1019
Db	623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
QY	1020	TGCCCTTTTACAAATGTGGCTATCTGCTTGATGCCTTTATATGATCATGTATACAATC	1079
Db	643	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAl	663
QY	1080	TAAAGAGCTTTCATCTTCTCGCCCACTTACAGGCCCTTTCTGTGTAAACAATATCTGAA	1139
Db	663	aLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuThrLysGlnTyrLeuAs	683
QY	1140	CTTTTACCCTGTCGGCGCAACGGTCGGTCTCTGCAAGTGTTCCTCACGCAACCCC	1199
Db	683	nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	703
QY	1200	CACTGATGGGGCTTGGCCATPAGCCATCAGCGCATGGCTGGAACTTTCTGGCTCTCTCT	1259
Db	703	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValSerProLe	723
QY	1260	GCCGATCCATACTCGGGAACCTCTAGACGCTTGTTCCTCGAGCCGGTCTGGAGCAAA	1319
Db	723	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	743
QY	1320	ACTTATCGGAACGACAACTCTGTTGTCTCTCTCGGAATATACACTCTCTTTCATGGCT	1379
Db	743	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	763
QY	1380	GCTAGGGTGTCTGCCAATCTGGATCTCTCGCGGACGCTCTTTGTCTACGTCGGTCCGC	1439
Db	763	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	783
QY	1440	GCTGAATCCCGGAGACGACCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
Db	783	aLeuAsnProAlaAspAspProSerArgLysArgLeuGlyLeuTyrArgProLeuLeuAr	803
QY	1500	TCTGCGTTCGGCGGACCAACGGGGCGCACTCTCTTTTACGCGGTCTCCCGTATGTGCC	1559
Db	803	gLeuProTyrArgProThrThrGlyArgThrSerLeuTyrAlaAspSerProSerValPr	823
QY	1560	TTCTCATCTGCGGACCGGTGTCACCTTCGTTTCACCTCTGCACGCTGCATGGAGACCAC	1619
Db	823	oSerArgLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	843
QY	1620	G 1620	
Db	843	O 843	

Search completed: December 2, 2005, 00:05:48
Job time : 152.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, (using frame_plus n2p model)

Run on: December 1, 2005, 23:14:00 ; Search time 464.5 Seconds
(without alignments)
6082.251 Million cell updates/sec

Title: US-10-761-006A-1
Perfect score: 5907
Sequence: 1 CTCACACATTCACACAG.....CCTCAGGCCAGCAGTGGA 3215

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cpn2.1/USPTO spool.p/US10761006/runat_01122005_113952_26862/app_query.fasta_1.3399
-DB=A_Geneseq -QFW=fastan -SUPPLX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -ALIGN=15 -MODE=LOCAL
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CNC 1.1.746 @runat_01122005_113952_26862 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 21:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2927	49.6	843	3 AAY54044	Aay54044 Amino aci
2	2810	47.6	540	9 ADX40768	Adx40768 HBV polym
3	2810	47.6	540	9 ADX40767	Adx40767 HBV polym
4	2807	47.5	843	9 ADX40772	Adx40772 HBV polym
5	2789	47.2	843	9 ADX40776	Adx40776 HBV polym
6	2778	47.0	845	9 ADX40777	Adx40777 HBV polym
7	2759	46.7	845	2 AAR70065	Aar70065 Hepatitis
8	2750	46.6	843	9 ADX40775	Adx40775 HBV polym
9	2737.5	46.3	539	7 ADM68212	Adm68212 GRP94 pro

10	2734.5	46.3	842	9	ADX40771	Adx40771 HBV polym
11	2688	45.5	825	9	ADX40774	Adx40774 HBV polym
12	2678.5	45.3	827	9	ADX40773	Adx40773 HBV polym
13	2675	45.3	842	4	AAB04708	Aae04708 Hepatitis
14	2675	45.3	842	4	AAG66918	Aag66918 HBV pres1
15	2654	44.9	832	9	ADX40761	Adx40761 HBV polym
16	2651	44.9	832	9	ADX40762	Adx40762 HBV polym
17	2643	44.7	843	9	ADX40760	Adx40760 HBV polym
18	2642	44.7	832	9	ADX40763	Adx40763 HBV polym
19	2628	44.5	832	9	ADX40765	Adx40765 HBV polym
20	2608	44.2	843	3	RAY44348	Aay44348 Human hep
21	2608	44.2	845	9	ADX40758	Adx40758 HBV polym
22	2607	44.1	832	9	ADX40764	Adx40764 HBV polym
23	2607	44.1	845	9	ADX40759	Adx40759 HBV polym
24	2582	43.7	843	9	ADX40766	Adx40766 HBV polym
25	1829	31.0	373	9	ADW42965	Adw42965 ADV (ade)
26	1829	31.0	373	9	ADW42948	Adw42948 ADV (ade)
27	1586	26.8	344	9	AEA08851	Aea08851 Hepatitis
28	1557	26.4	344	9	AEA08850	Aea08850 Hepatitis
29	1552	26.3	303	9	ADX40769	Adx40769 HBV polym
30	1552	26.3	303	9	ADX40770	Adx40770 HBV polym
31	1496	25.3	400	3	RAY54045	Aay54045 Amino aci
32	1479	25.0	309	9	AEA08832	Aea08832 Hepatitis
33	1465	24.8	281	1	AAP60163	Aap60163 Subtype a
34	1465	24.8	281	1	AAP60560	Aap60560 Hepatitis
35	1465	24.8	281	1	AAP60617	Aap60617 C-termina
36	1462	24.8	307	9	AEA08833	Aea08833 Hepatitis
37	1461	24.7	281	1	AAP70294	Aap70294 Subtype a
38	1459	24.7	400	2	AAR93801	Aar93801 Hepatitis
39	1459	24.7	400	6	ABR55863	Abt55863 HBV S-pro
40	1459	24.7	400	8	ADN49718	Adn49718 Hepatitis
41	1459	24.7	400	8	ADU74394	Adu74394 HBV S-pro
42	1457	24.7	348	7	ADG76935	Adg76935 Hepatitis
43	1456	24.6	281	2	AAR62870	Aar62870 Hepatitis
44	1456	24.6	281	4	AAM48419	Aam48419 Protein #
45	1452	24.6	389	1	AAP60794	Aap60794 Adr-type

ALIGNMENTS

RESULT 1

RAY54044
ID AAY54044 standard; protein; 843 AA.

XX AC AAY54044;
XX AC
XX 27-MAR-2000 (first entry)
XX
XX DE Amino acid sequence of a HBV DNA polymerase protein.

XX KW HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase;
XX KW large surface antigen; core protein; transactivating X protein;
XX KW hepatitis vaccine; HBV infection; hepatocellular carcinoma.

XX OS Hepatitis B virus.

XX FH Key Location/Qualifiers
XX FT Misc-difference 194 /note= "encoded by GAG"
XX FT Misc-difference 195 /note= "encoded by CAA"
XX FT

XX PN WO9966048-A1.

XX XX 23-DEC-1999.

XX PF 19-JUN-1998; 98WO-SG0000046.

XX PR 19-JUN-1998; 98WO-SG0000046.

XX PA (GOVE-) GOVERNMENT REPUBLIC SINGAPORE.

XX XX Oon CJ, Lim GK, Zhao Y, Chen WN;

```
XX WPI; 2000-106104/09.
DR N-ESDB; AAZ37088.
XX
XX New isolated hepatitis B virus strain, useful for, e.g. treatment of
PT hepatitis infection.
XX
PS Disclosure; Page 36-39; 68pp; English.
XX
CC The present sequence is encoded by the genome of an isolated strain of
CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
CC antigen-'S'-133 Oon strain (Met to Thr). The viral genome is deposited as
CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
CC large surface antigen, a core protein, and a transactivating X protein.
CC The large surface antigen differs from the wild type sequence in that it
CC contains a Thr at position 133 of the wild type sequence instead of a
CC Met. The proteins are used to produce antibodies. The proteins,
CC polynucleotide and antibodies can be used for detecting the novel HBV
CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
CC HBV novel strain polypeptides can be used to identify compounds for
CC treating or preventing HBV infection or hepatocellular carcinoma
XX
XX Sequence 843 AA;
SQ
Alignment Scores:
Pred. No.: 9.31e-224 Length: 843
Score: 2927.00 Matches: 540
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.55% Indels: 0
DB: 3 Gaps: 0
US-10-761-006A-1 (1-3215) x AAY54044 (1-843)
QY 1 CTCACACATTCACCAAGCTCTGTAGATCCAGAGGTGAGGGGCTATATTTCTCTGC 60
DB LeuHisAsnIleProProSerSerAlaArgSerGlnGlyGluGlyProIlePheSerCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAAACCCCTGTTCCGACTCTCCCTCCCATATCGTC 120
DB TrpTrpIleuGlnPheArgAsnSerLysProCysSerAspIleVal 343
QY 121 AATCTTCTCGAGGACTCGGGACCTCGCACCGAAATCGAGAACACATCAGGATTCCT 180
DB AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTGTGTACAGCGGGGTTTCTCGTTGACAAAGATCCTCACAATACC 240
DB ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTGCTGGACTCTCTCAATTTCTAGGGGGAGCACCCAGCTGTTC 300
DB AlaGluSerArgLeuTrpTrpThrSerLeuAsnPheLeuGlyAlaProThrCysSer 403
QY 301 TGGCCAAATTCGAGTCCCAACCTCCAATCACTACCAACCTTTGTCTCCCAATTG 360
DB TrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeu 423
QY 361 TCCTGGCTATCGTGGATGTCTGCGCGGTTTATCATATTCCTCTCATCTCTGCTGCT 420
DB SerTrpLeuSerLeuAspValSerAlaAlaPheThrHisIleProLeuHisProAlaAla 443
QY 421 ATGCTCATCTCTGTGTGGTCTCTCTGGACTACCAAGGTATGTTGCCGTTGTCTCT 480
DB MetProHisLeuLeuValGlySerSerGlyLeuProArgIleValAlaArgLeuSerSer 463
QY 481 ACTTCCAGGAACATCAACACAGCACGGGGCCATGCAAGACCTGACGACTCTCTGCTCA 540
DB ThrSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSer 483
QY 541 AGGAACTCTACGTTTCCCTCTTGTCTGTACAAAACCTTCGGACGGAACTGCATTTG 600
484 ArgLysLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeu 503
601 TATTCCCATCCCATCATCTGGGCTTTCCGAAATTCCTATGGAGTGGGCGCTCAGTCCG 660
504 TyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPro 523
661 TTTCTCTCTGGCTCAGTTTACTAGTCCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCAC 720
524 PheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHis 543
721 TGTTGGCTTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 780
544 CysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLeu 563
781 GAGTCCCTTTTACTCTATTTACCAATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 840
564 GluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPro 583
841 AATAAAACCAAGCTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTTGGAAGTTGG 900
584 AsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrp 603
901 GGTACTTTACCGCAGGACATATTTGACTAAACTCAAGCAATGTTTTTCGAAACTGCCT 960
604 GlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPro 623
961 GTAAATAGACTATTGATTCGAAAGATATGTCGAAAGATTTGGTCTTTTGGGCTTTGCT 1020
624 ValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAla 643
1021 GCCCTTTTACACAATGTGGCTATCTCGCTTGCCTTATGCTTTATATGATGATGATGATG 1080
644 AlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSer 663
1081 AGCAGGCTTTTCATTTCTCGCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAAC 1140
664 LysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAsn 683
1141 CTTTACCCGCTGTCGGGCAACGCTCCGCTCTCGCAAGTGTTCGTGACGCAACCCCTC 1200
684 LeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPro 703
1201 ACTGGATGGGCTTGGCCATAGGCCATCAGCGCATGCTGGAACCTTTCTGGCTCTCTCTG 1260
704 ThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetAlaGlyThrPheLeuAlaProLeu 723
1261 CGCATCCTACTCGGGAACCTCTAGCAGCTTTGCTGCTCGCAGCCGCTCTGGAGCAAAA 1320
724 ProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 743
1321 CTTATCGGAACCGACAACTCTGTTGCTCTCTCGGAAATACACCTCTCTTTCATGGCTG 1380
744 LeuIleGlyThrAspAsnSerValValLeuSerArgLysIleThrSerPheProTrpLeu 763
1381 CTAGGGTGTCTGCCAATCTGGATCCTCGCGGGAACGCTCTTGTCTACGTCCTCGGCG 1440
764 LeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAla 783
1441 CTGAATCCCGCGGACGACCCGCTCTCGGGGCGGCTTTGGGCTCTACCGTCCCTCTTTCAT 1500
784 LeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHis 803
1501 CTCCGCTTCCGGCGGACCGGCGGCGCTCTCTTTCAGCGGCTCTCCCGTATGTCCT 1560
804 LeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProTyrValPro 823
1561 TCTCATCTGCGGACCGTGTGCTCTGCTTACCTCTGCTGACCGTCTGATGAGACACCG 1620
824 SerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
RESULT 2
ADX40768
```


ID ADX40768 standard; protein; 540 AA.
 XX AC ADX40768;
 XX DT 21-APR-2005 (first entry)
 XX DE HBV polymerase protein #11.
 XX KW Immune stimulation; polymerase; enzyme.
 XX OS Hepatitis B virus.
 XX PN WO2005012502-A2.
 XX PD 10-FEB-2005.
 XX PF 29-MAR-2004; 2004WO-US009510.
 XX PR 28-MAR-2003; 2003US-0458026P.
 XX PA (EPIM-) EPIMUNE INC.
 XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 XX WIPI; 2005-132661/14.
 XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX PS Disclosure; Page 380-385; 458pp; English.
 XX SS The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX SQ Sequence 540 AA;
 Alignment Scores:
 Pred. No.: 1.68e-214 Length: 540
 Score: 2810.00 Matches: 527
 Percent Similarity: 98.34% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 8
 Query Match: 47.57% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006A-1 (1-3215) x ADX40768 (1-540)
 QY 1 CTCACAAATCCACAGCTCTGATCCAGGGGTGAGGGGCTATATTTCTGTC 60
 Db 1 LeuHieAsnIleProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCys 20
 QY 61 TGTGTGCTCCAGTTCGGACAGTAACCTGTTCGACTACTGCTCTCCCATATCGTC 120
 Db 21 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspPyrCysLeuThrHisIleVal 40
 QY 121 AATCTTCTCGAGACTGGGACCTCGACCGAACATCGAGAACACAACTCAGGATTCCT 180
 Db 41 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
 QY 181 AGACCCCTGCTCGTGTACAGCGGGGTTTTCTCGTTGACAAAGATCTCAATATACC 240
 Db 61 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 80
 QY 241 GCAGACTAGACTCTG-GTGGACTTCTCAATTTCTAGGGGGAGACCCACAGGTTC 299
 Db 81 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100

QY 300 CTGCGCAAAATTCGAGTCCCAACCTCCATCACTCACCAACCTCTTGTCTCTCAATTT 359
 Db 100 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe 120
 QY 360 GTCTCTGGCTATCGTGGATGTGTCTGCGGCGTTTTATCATATTCCTCTTCATCTGCTGC 419
 Db 120 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyHisIleProLeuHisProAlaAl 140
 QY 420 TATGCTCATCTTCTGTGTGGTCTCTTGGAGTACCAAGGTATGTGGCCGTTGTCTCTC 479
 Db 140 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyValAlaArgLeuSerSe 160
 QY 480 TACTTCCAGGAACATCAACACCGACGCGGCCCATGCAAGACCTCGACGACTCTCTGCTC 539
 Db 160 rThrSerArgAsnIleAsnTyGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
 QY 540 AAGGAAACTCTACGTTTCCCTCTCTGTGTGTACAAAACCTTCGGACGGAACATGCACTT 599
 Db 180 rArgAsnLeuTyValSerLeuLeuLeuLeuTyHisThrPheGlyArgGlyLeuHisLe 200
 QY 600 GTATTCCCATCCCATCATCTCTGGGCTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
 Db 200 uTy-SerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
 QY 660 GTTCTCTCTGGCTCAGTTTACTAGTCCCATTTGTTCAGTGGTTCGTAGGGCTTCCCCCA 719
 Db 220 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 240
 QY 720 CTGTTTGGCTTTCAGTATATGATGTGTGTATTTGGGGCGAAGTCTGTACAACATCT 779
 Db 240 eCysLeuAlaPheSerTyMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
 QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
 Db 260 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
 QY 840 TAAATAAACCAACCTGGGCTACTCCCTTAACCTCATGGATATGTAAATTTGGAGTTG 899
 Db 280 oAsnLysThrLysArgTrpGlyTy-SerLeuAsnPheMetGlyTyValIleGlySerTr 300
 QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAACCTCAAGCAATGTTTTCGAAAACTGCC 959
 Db 300 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
 QY 960 TGTAAATAGACCTATTGTATTTGAAAGTATGTCAAGAATTTGGGTCTTTTGGGCTTTC 1019
 Db 320 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
 QY 1020 TGCCCTTTTACAAATGTGGCTATCTCTGCTTGTATGCTTATATGATGTATATCAATC 1079
 Db 340 aAlaProPheThrGlnCysGlyTy-ProAlaLeuMetProLeuTyAlaCysIleGlnSe 360
 QY 1080 TAAGCAGGCTTTTCACTTTCTCGCCCACTTACAAAGGCTTTCTGTGTAAACAATATCTGAA 1139
 Db 360 rLysGlnAlaPheThrPheSerProThrTyLysAlaPheLeuCysLysGlnTyLysAs 380
 QY 1140 CCTTTACCCCGTTCGCGCAACGGTCCGGTCTCTGCCAAGTGTTCGTGACGCAACCCC 1199
 Db 380 nLeuTyProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 QY 1200 CACTGTAGTGGGCTGGCCATAGGCCATCAGGCATCGCTGGACCTTTCTGCTCTCTCT 1259
 Db 400 oThrGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 420
 QY 1260 GCGCATCCATCTCGGGAACCTCTAGCAGCTGTTTTGTCTCGCAGCGGTCTCGAGCAAA 1319
 Db 420 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLy 440
 QY 1320 ACTTATCGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTCTTCCATGGCT 1379
 Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyThrSerPheProTrpLe 460

QY 1380 GCTAGGGTGTGCTGCTCCAACTGGATCTCGCGGGAGCGTCTTTGTCTACGTCCCGTGGC 1439
Db ||||| uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
QY 1440 GCTCAATCCCGCGGACGACCGTCTCGGGGCGGTTGGGGCTCTACCGTCCCTCTTCA 1499
Db ||||| aLeuAsnProAlaAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
QY 1500 TCTGCGGTTCCGGCGGACACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
Db ||||| sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
QY 1560 TTCTCATCTGCGGACCGGTGCGACTTCGCTTACCTCTGCACTGCGCATGGAGACCA 1619
Db ||||| oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
QY 1620 G 1620
Db 540 o 540

RESULT 3

ID ADX40767 standard; protein; 540 AA.

AC ADX40767;

DT 21-APR-2005 (first entry)

DE HBV polymerase protein #10.

KW Immune stimulation; polymerase; enzyme.

OS Hepatitis B virus.

PN WO2005012502-A2.

PD 10-FEB-2005.

PF 29-MAR-2004; 2004WO-US009510.

PR 28-MAR-2003; 2003US-0458026P.

FA (EPIM-) EPIMMUNE INC.

PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

DR WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

XX Disclosure; Page 380-385; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.

XX Sequence 540 AA;

Alignment Scores:

Pred. No.: 1.68e-214 Length: 540
Score: 2810.00 Matches: 527
Percent Similarity: 98.34% Conservative: 5
Best Local Similarity: 97.41% Mismatches: 8
Query Match: 47.57% Indels: 2
DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40767 (1-540)

QY 1 CTCCACACATTCACCAAGCTCTGCTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGC 60
Db ||||| LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 20
QY 61 TGGTGGCTCCAGTTCCCGAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCTCTC 120
Db ||||| TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 40
QY 121 AATCTTCTCGAGGACTGGGACCTGCAACCAATCGAGAACACAAATCAGATTCTCT 180
Db ||||| AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 60
QY 181 AGGACCCCTGCTGTTTACAGGGCGGGTTTTCTCGTTGACAAGATCTCTCACAAATACC 240
Db ||||| ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr 80
QY 241 GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTTCTAGGGGAGCACCACCTGTTC 299
Db ||||| ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-Se 100
QY 300 CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACCACTCTTGTCTCTCAATTT 359
Db ||||| rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 120
QY 360 GTCTGGCTATCGCTGGATGCTCTGGGGGTTTTATCATATTTCTTCTATCTGCTGTC 419
Db ||||| uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 140
QY 420 TATGCTCATCTCTGTTGTTGTTCTCTGGACTACCAAGGTATGTTGCCCTTTGTCTCTC 479
Db ||||| aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 160
QY 480 TACTTCCAGGAACATCAACACACAGCAGCGGGCCATGCAAGACCTGCACTGCTGCTC 539
Db ||||| rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 180
QY 540 AAGGAAACTCTACGTTTCCCTCTGTTGCTGTACAAAACCTTCGGACGGAACCTGCAC 599
Db ||||| rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 200
QY 600 GTATTTCGCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGAGTGGGCCCTCAGTCC 659
Db ||||| uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 220
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTGAGTGGTTCGTAGGGCTTTTCCCCCA 719
Db ||||| oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi 240
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db ||||| sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 260
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTGTTGTTGGTATACATTTTAAACCC 839
Db ||||| uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 280
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTAATTGGAAGTTG 899
Db ||||| oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 300
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACCTGCC 959
Db ||||| pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 320
QY 960 TGTAAATAGACCTATTGATTGGAAGTATGTCAAAGAAATGTTGGGTCTTTTGGGCTTTC 1019
Db ||||| oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 340
QY 1020 TGCCCTTTTACAAATGTGGCTATCTGCTTGTATGCTTGTATATGCTTGTATATGCTAT 1079
Db |||||

Db 340 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 360
 Qy 1080 TAAGCAGGCTTCTACTTTCTCCCAACTTACAGGCGCTTCTGTGTAACAAATATCTGAA 1139
 Db 360 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysIleGlnTyrLeuAs 380
 Qy 1140 CTTTACCCCTTGGCCGCAACGCTCCGCTCTCTGCAAGTGTGCTGACGCAACCCC 1199
 Db 380 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 400
 Qy 1200 CACTGATGGGCTTGGCCATAGGCAATCAGCGCATGGCTGGAACCTTCTGGCTCTCT 1259
 Db 400 oThrGlyTyrGlyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe 420
 Qy 1260 GCGATCCATCTGCGGAACCTCTAGCAGCTTGTGTCGCGCGCTCTGAGCAAA 1319
 Db 420 uProIleHisThrAlaGluLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys 440
 Qy 1320 ACTTATCGAACCGCAACTCTGTTGCTCTCTCGGAATACACCTCTTCCATGGCT 1379
 Db 440 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProtrPle 460
 Qy 1380 GCTAGGCTGTGCTCAACTGATCTCTGCGGAGCGTCTTGTCTACGTCCTCCGCGC 1439
 Db 460 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 480
 Qy 1440 GCTGAATCCCGGAGCAGCCCTCTCGGGCGCTTGGGGCTCTACCGTCCCTTCTCA 1499
 Db 480 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 500
 Qy 1500 TCTGCGCTTCCGGCCGACACCGGGCGACCTCTCTTACGGGCTCTCCCGTATGTGCC 1559
 Db 500 sLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 520
 Qy 1560 TTCTCATCTGCGGACCGGTGTCACCTTCTGCTGACGTCGATGAGAGCACCC 1619
 Db 520 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 540
 Qy 1620 G 1620
 Db 540 o 540
 RESULT 4
 ID ADX40772
 XX ADX40772 standard; protein; 843 AA.
 AC ADX40772;
 XX 21-APR-2005 (first entry)
 DT HBV polymerase protein #15.
 DE Immune stimulation; polymerase; enzyme.
 KW Hepatitis B virus.
 OS WO2005012502-A2.
 XX 10-FEB-2005.
 XX 29-MAR-2004; 2004WO-US009510.
 XX 28-MAR-2003; 2003US-0458026P.
 XX (EPIN-) EPIMUNE INC.
 PA Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI WPI; 2005-132661/14.
 DR Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class

PT I binding motif.
 XX Disclosure; Page 380-385; 459pp; English.
 CC The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX Sequence 843 AA;
 SQ
 Alignment Scores:
 Pred. No.: 3.31e-214 Length: 843
 Score: 2807.00 Matches: 528
 Percent Similarity: 98.33% Conservative: 3
 Best Local Similarity: 97.78% Mismatches: 8
 Query Match: 47.52% Indels: 2
 DB: 9 Gaps: 0
 US-10-761-006A-1 (1-3215) x ADX40772 (1-843)
 Qy 4 CACACATTCACCAAGCTCTGTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGCTGG 63
 Db 305 HisAsnIleProProSerSerAlaArgSerGlnSerGlnGlyProIlePheSerCysTrp 324
 Qy 64 TGGCTCCAGTTCGGAACAGTAAACCTGTTCGCACTACTGCTCTCCCATATCGTCAAT 123
 Db 325 TrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleValAsn 344
 Qy 124 CTTCTCAGGACTGGGACCCCTGCAACGAAATGAGAACACAAATCAGAGTTCTTAGG 183
 Db 345 LeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArg 364
 Qy 184 ACCCTGCTGCTGTGTACAGGGGGGTTTTCTGCTTGACAGAACTCCACAAATCCGCA 243
 Db 365 ThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThrThr 384
 Qy 244 GAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACCCACGTTGCTCTG 302
 Db 385 GluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisVal-SerTr 404
 Qy 303 GCCAAATTCGAGTCCCAACTCCAATCACTCAACCACTCTGTCTCTCCAAATTTGTC 362
 Db 404 pProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSe 424
 Qy 363 CTGGCTATCGCTGGATGTCTGCGGGTTTTATCATATTTCTCTTCATCTGCTGCTAT 422
 Db 424 rTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAlaMe 444
 Qy 423 GCCTCATCTTCTGTTGTTGTTCTCTGGACTACCAAGGTATGTTGCCGTTTGTCTCTAC 482
 Db 444 tProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerTh 464
 Qy 483 TTCCAGGAACATCAACACCAAGCGGGCCATGCAAGACCTGACAGCTCTCTGCTCAAG 542
 Db 464 rSerArgAsnIleAsnHisGlnHisGlyAlaMetGlnAspLeuHisAspSerCysSerAr 484
 Qy 543 GAAACTCTAGCTTTCCTCTTGTGTGTACAAACCTTCGGACGGAACTGACCTGTA 602
 Db 484 gAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTy 504
 Qy 603 TTCCCATCCCATCTCTGCGGCTTTTCGCAAGATTCCTATGAGAGTGGGCTCAGTCCGTT 662
 Db 504 rSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPh 524
 Qy 663 TCTCTGCTGCTCAGTTTACTAGTGCCATTTGTTTTCAGTGGTTTCGTAGGGCTTCCCCC 722
 Db 524 eLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCy 544

```
Qy 723 TTTGGCTTTCAGTTATATATGATGATGTTGGGGCGAAGTCTGTACAAATCTTGA 782
Db |||||
Qy 744 sLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGl 564
Db |||||
Qy 783 GTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCCCTAA 842
Db |||||
Qy 564 uSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAs 584
Db |||||
Qy 843 TAAACCAACAGTTGGGCTACTCCCTTAACTTCATGGGATATGTAAATTGGAAGTTGGG 902
Db |||||
Qy 584 nLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGl 604
Db |||||
Qy 903 TACTTTACCGCAGGACATATTGTACTTAAACTCAAGCAATGTTTTCGAAACTGCCTGT 962
Db |||||
Qy 604 yThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuProVa 624
Db |||||
Qy 963 AAATAGACCTATTGATTGGAAGATGTCAAGAATGTGGGTCTTTTGGGCTTTGGTC 1022
Db |||||
Qy 624 lAsnSerProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAl 644
Db |||||
Qy 1023 CCCTTTTACAAATGTGGCTATCTCGCTTCATGCTCTTATATGATCATGTATACAACTAA 1082
Db |||||
Qy 644 aProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSerLy 664
Db |||||
Qy 1083 GCAGGCTTTCATCTTCTCGCAACTTACAGGCTTCTGTGTAAACATATCTGAACCT 1142
Db |||||
Qy 664 sGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuLeuGlyGlnTyrLeuAsnLe 684
Db |||||
Qy 1143 TTACCCCTGTCGCCGCAACGTCGGGTCTCTGCCAAGTGTTCGTACGCAACCCCAAC 1202
Db |||||
Qy 684 uTyrProValAlaArgGlnArgSerGlyLeuLysGlnValPheAlaAspAlaThrProTh 704
Db |||||
Qy 1203 TGAATGGGGCTTGCCATAGCCATCAGCGCATGCTGGAACTTCTTGGCTCTCTGCC 1262
Db |||||
Qy 704 rGlyTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuPr 724
Db |||||
Qy 1263 GATCCATCTCGGAACTCTAGCAGCTTGTCTCGAGCGGCTGTGAGCAAACT 1322
Db |||||
Qy 724 oIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLe 744
Db |||||
Qy 1323 TATCGGAACCACTCTGTGTCTCTCTCGGAATATACACCTCTTCCATGCGTCT 1382
Db |||||
Qy 744 uIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLe 764
Db |||||
Qy 1383 AGGTGTGTCGCAACTGGATCTCGCGGACGTCTCTTGTCTACGTCCCGTGGCGCT 1442
Db |||||
Qy 764 uGlyCysAlaAlaAsnTrpIleLeuA-gGlyThrSerPheValTyrValProSerAlaLe 784
Db |||||
Qy 1443 GAATCCGCGGACGACCGCTCTCGGGCCGTTTGGGCTCTACCGTCCCTTCTTCATCT 1502
Db |||||
Qy 784 uAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuLe 804
Db |||||
Qy 1503 GCCGTTCGCGCGACACCGGCGCACCTCTCTTTTACGCGTCTCTCCGATGTGCTTC 1562
Db |||||
Qy 804 uProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSe 824
Db |||||
Qy 1563 TCATCTCGCGACCGTGTGCACTTCGCTTCACTCTGACGTCTGACGTGAGACACCG 1620
Db |||||
Qy 824 rHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 843
Db |||||
RESULT 5
ADX40776
ID - ADX40776 standard; protein; 843 AA.
XX AC
AC AC
XX AC
XX AC
DT 21-APR-2005 (first entry)
XX DE
DE HBV polymerase protein #19.
XX KW
KW Immune stimulation; polymerase; enzyme.
XX XX
XX Hepatitis B virus.
```

XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 380-385; 450pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX SQ Sequence 843 AA;

Alignment Scores:
Pred. No.: 8.98e-213 Length: 843
Score: 2789.00 Matches: 523
Percent Similarity: 98.15% Conservative: 8
Best Local Similarity: 96.67% Mismatches: 9
Query Match: 47.22% Indels: 2
DB: Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40776 (1-843)

Qy 1 CTCACCAACATTCACCAAGCTCTGTAGATCCAGGGTGAGGGGCTATATTTCTGTC 60
Db 304 LeuHisAsnIleProProSerSerAlaArgSerGlnSerGluGlyProIlePheSerCys 323
Qy 61 TGTGGCTCCAGTTCGGAAACAGTAACCCCTGTTCGACTACTGCTCTCCCATATCGTC 120
Db 324 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
Qy 121 AATCTTCTCGAGGACTGGGACCTCGACCAATGAGAGACACATCATCAGATTCCT 180
Db 344 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
Qy 181 AGGACCCCTGCTCGTGTTCACAGCGGGGTTTTCTCGTTCACAGAATCTCTCAAAATACC 240
Db 364 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
Qy 241 GCAGAGCTAGACTCTG- GTGGACTTCTCTCAATTTCTAGGGGGAGACCCACGTGTC 299
Db 384 ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrHisIleVal-Se 403
Qy 300 CTGGCCAAAATTCGACGTCCCAACCTCCAATCACTCACCAACCTCTGTGCTCCAATTT 359
Db 403 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
Qy 360 GTCTGTGCTATCGCTGGATGTGTGCGGGGTTTTATCATATTCTCTTCATCTGCTGTC 419
Db 423 uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
Qy 420 TATGCTCATCTTCTGTGGTCTTCTGTGACTACCAAGGTATGTTGCCGCTTGTGCTC 479
Db |||||

```
Db 443 aMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACCAACAGCGGCCCATGCAAGACCTGCACGACTCTCTGCTC 539
Db 463 rThrSerArgAenIleAenTyrGlnHisGlyThrMetGlnAepLeuHisAepSerCysSe 483
QY 540 AAGGAACCTACAGTTTCCCTCTTCTGCTGTACAAAACCTTCGACGGAACCTGCACTT 599
Db 483 rArgAenLeuTyrValSerLeuLeuValTyrLysThrPheGlyArgLeuHisLe 503
QY 600 GTATTCCCATCCATCATCTCGGCTTCGCAAGATTCTATGGAGTGGGCTCAGTCC 659
Db 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCGCTCAGTTTACTAGTCCCATTTGTTTCAAGTGTTCAGTGGTTCAGGGCTTCC 719
Db 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 543
QY 720 CTGTTTGGCTTTCAGTTATATGATGATGTGTATTTGGGGCGAGTCTGTACACATCT 779
Db 543 sCysLeuAlaPheSerTyrMetAspAepValValLeuGlyAlaLysSerValGlnHisLe 563
QY 780 TCAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db 563 uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTCATGGGATATGTAATTGGAAGTTG 899
Db 583 oAenLysThrLysArgTgPlyTyrSerLeuAenPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTATAAACTCAAGCAATGTTTTCGAAAACCTGCC 959
Db 603 pGlyThrLeuProGlnGluHisIleValLeuLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATAGACTATTGATTGGAAGTATGTAAGAAATTTGGGTCTTTGGCTTTCG 1019
Db 623 oValAenArgProIleAspIrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCCCTTTTACACATATGGCTATCTCTGCTTGATGCTTCTTATATGATGATATCAATC 1079
Db 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCATCTTCGCCAATCTTCAAGGCTTCTGTGTAAACATATCTGAA 1139
Db 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysGlnGlnTyrLeuHi 683
QY 1140 CTTTACCCGTTGCGCGCAAGGTCGCTCTCTCCAAAGTGTGCTGAGCAACCC 1199
Db 683 sLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheGlyAspAlaThrPr 703
QY 1200 CACTGATGGGCTTCGCCATAGCCATCAGCCATGGCTGGAACCTTTCTGCTCTCT 1259
Db 703 oThrGlyTgPlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCGATCCATATCGCGAACTCTAGCAGTTGTTTGTCTGCGAGCGGCTCGGACAA 1319
Db 723 uProIleHisThrAlaGluLeuLeuAlaCysPheAlaArgSerArgSerGlyAlaLys 743
QY 1320 ACTTATCGGAACCGAACCTCTGCTCTCTCGGAATACACCTCTTTCATGGCT 1379
Db 743 sLeuIleGlyThrAspAenSerValValLeuSerArgLysTyrThrSerPheProIrpLe 763
QY 1380 GCTAGGCTGTGTCGCAACTGGAATCTTCGCGGAGCTCTTGTCTAGCTCCGCTCGGC 1439
Db 763 uLeuGlyCysAlaAenTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGAGCAGCCGCTCTCGGGCCGTTTGGGCTCTACCGTCCCTCTTCA 1499
Db 783 aLeuAenProAlaAspAepProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHi 803
QY 1500 TCTGCCGTTCCGGCGACGAGGGGCGACCTCTTTTACGGGCTCTCCCGTATGTC 1559
Db 803 sLeuProPheGlnProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
```

```
QY 1560 TTCTCATCTCCGACCGCTGTGCACTTCGCTTCACTCTGACCTCGCATCGAGACCACC 1619
Db 823 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaIleTrpArgProPr 843
QY 1620 G 1620
Db 843 o 843
RESULT 6
ADX40777
ID ADX40777 standard; protein; 845 AA.
XX ADX40777;
XX 21-APR-2005 (first entry)
XX HBV polymerase protein #20.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis B virus.
XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 380-385; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX Sequence 845 AA;
Alignment Scores:
Pred. No.: 6,74e-212 Length: 845
Score: 2778.00 Matches: 521
Percent Similarity: 97.41% Conservative: 6
Best Local Similarity: 96.30% Mismatches: 13
Query Match: 47.03% Indels: 2
DB: Gaps: 0
US-10-761-006A-1 (1-3215) x ADX40777 (1-845)
```

```
QY 1 CTCCCAACATTCACCAAGCTCTGTAGATCCAGGTGAGGGCTATATTTCTCTGC 60
Db 306 LeuHisAsnIleProProSerAlaArgSerGlnSerGluGlyPro***PheSerCys 325
QY 61 TGTGTGCTCCAGTTCGGAACAGTAAACCTGTGTCCGACTACTGCTCTCCCATATGTC 120
Db 326 TrpTrpLeuGlnPheArgAenSerLysProCysSerAspTyrCysLeuSerHisLeVal 345
```

QY 121 AATCTTCTCGAGGACTGGGACCTTCGACCGAACATGAGAGAACACAAATCAGGATTCCT 180
Db |||||
346 AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisIleArgIlePro 365
QY 181 AGGACCCCTGCTCGTGTACAGGGGGGTTTCTCGTTGACAGAAATCCTCACAATACC 240
Db |||||
366 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 385
QY 241 GCAGAGTCTAGACTCTG- GTGACTTCTCTCAATTTCTAGGGGAGCACCCAGCTGTTC 299
Db |||||
386 ThrGluSerArgLeuValIleAspPheSerGlnPheSerArgGlySerThrArgVal-Se 405
QY 300 CTGGCCAAAATCGAGTCCCAACCTCCAATCACTCAACACCTCTTGCTCCTCAATTT 359
Db |||||
405 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 425
QY 360 GTCTCGGTATCTCGTGGATGTCTCGCGGCTTTTATCATATTTCTTCATCTCTGCTGC 419
Db |||||
425 uSerTrpLeuSerLeuAspValSerAlaIlePheThrHisIleProLeuHisProAlaAl 445
QY 420 TATGCTCATCTCTTGTGTGTTCTCTGACTACCAAGGTATGTGCGGCTTTGTCTCTC 479
Db |||||
445 aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe 465
QY 480 TACTTCCAGGAACATCAACACCCAGCAGCGGGCCATGCAACCTCGACGACTCTGCTC 539
Db |||||
465 rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe 485
QY 540 AAGGAAACTCTAGCTTCCCTCTGTGCTGTACAAACCTTCGGAGCAACTGCACCTT 599
Db |||||
485 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe 505
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATTCCTATGGAGTGGGCTCAGTCC 659
Db |||||
505 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 525
QY 660 GTTCTCTCGCTCAGTTTACTAGTGCATTTGTTAGTGTTCGTAGGGCTTTCCCCCA 719
Db |||||
525 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 545
QY 720 CTGTTTGGCTTTCAGTTATATGATGTGTGTATGGGGCGAAGTCTGTACAACTCT 779
Db |||||
545 sCysLeuAlaPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLe 565
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTCTTTTGTCTTTGGGTATACATTTAAACCC 839
Db |||||
565 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 585
QY 840 TAATAAAACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGTAATGGAAGTTG 899
Db |||||
585 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 605
QY 900 GGGTACTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACTGCC 959
Db |||||
605 pGlyThrLeuProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr 625
QY 960 TGTAATAGACTATTGATTGGAAAGTATGTCAAGAAATTCGGGTCTTTTGGGCTTTCG 1019
Db |||||
625 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 645
QY 1020 TGCCGCTTTTACAAATGTGCTATCTGCTTCCTGCTTCATGCTTTATATGATGATACAAATC 1079
Db |||||
645 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 665
QY 1080 TAAGCAGGCTTTCACTTTCTCGCAACTTACAAGGCTTCTGTGTAAACAAATATCTCAA 1139
Db |||||
665 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 685
QY 1140 CTTTACCCGCTTTCGGGCAACGCTCGGCTCTCTGCCAAGTGTGTTGCTGACGCAACCCC 1199
Db |||||
685 nLeuTyrProValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 705
QY 1200 CACTGGATGGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGGCTCTCT 1259

Db |||||
705 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 725
QY 1260 GCCGATCCATACTCGCGAACTCCTAGCAGCTTGTGTTGCTCGCAGCCGCTCGAGCAAA 1319
Db |||||
725 uProIleHisThrAlaGluLeuLeuAlaIleCysPheAlaArgSerArgSerGlyAlaLys 745
QY 1320 ACTTATCGGAACCACTCTGTTGTCTCTCTCGGAAATACACCTCTCTTCCATGGCT 1379
Db |||||
745 sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 765
QY 1380 GCTAGGTGTCTCCCAACTCGATCTCGCGGGGACGTCCTTGTCTACGTCCTCGGCTCGGC 1439
Db |||||
765 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 785
QY 1440 GCTCAATCCGCGACGACCGCTCTCGGGGCGGCTTTCGGGCTCTACCGTCCCTCTTCA 1499
Db |||||
785 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 805
QY 1500 TCTGCCGTTCCGGCGCACCGACCGCGCTCTCTTTTACGCGGTCTCCCCGTATGTGCC 1559
Db |||||
805 gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 825
QY 1560 TTCTCATCTGCGCGACCGTGTGCACTTCGCTTACCTCTGCACTGCGATGCGAGACACC 1619
Db |||||
825 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 845
QY 1620 G 1620
Db |||||
845 o 845
RESULT 7
AAR70065
ID AAR70065 standard; protein; 845 AA.
AC AAR70065;
XX 25-MAR-2003 (revised)
DT 06-OCT-1995 (first entry)
XX Hepatitis B virus polymerase protein.
XX Hepatitis B virus polymerase; cytotoxic T cell response; prophylactic;
XX vaccine; chronic; acute HBV infection; carrier.
XX Hepatitis B virus.
XX WO9503777-A1.
XX 09-FEB-1995.
XX 01-AUG-1994; 94WO-US008685.
XX 02-AUG-1993; 93US-00100870.
XX (SCRI) SCRIPPS RES INST.
XX Chisari FV;
XX WPI; 1995-082004/11.
XX New peptides inducing cytotoxic T lymphocytes to hepatitis B virus - are
XX regions of HB polymerase protein, for treating acute and chronic
XX infections.
XX Disclosure; Page 50-52; 85pp; English.
XX The amino acid sequence of the Hepatitis B virus (HBV) polymerase (HBpol)
XX protein. The sequence was used to generate a series of peptides (AAR70044
XX -59) which induce cytotoxic T cell (CTL) responses against cells infected
XX with HBV. The HBpol peptides can be used, prophylactically as vaccines,
XX together with, or conjugated to, epitopes from other HBV sequences that
XX elicit T cell responses to HBV (see AAR70060-64). The peptides can be

CC used, particularly ex vivo, to stimulate CTL cells. These cells can be
 CC reintroduced into patients who have chronic or acute HBV infections or
 CC are carriers, especially in treatments to prevent conversion from acute
 CC to chronic infections. (Updated on 25-MAR-2003 to correct PN field.)

XX
 SQ Sequence 845 AA;

Alignment Scores:

Pred. No.: 2,19e-210 Length: 845
 Score: 2759.00 Matches: 516
 Percent Similarity: 97.23% Conservative: 10
 Best Local Similarity: 95.38% Mismatches: 14
 Query Match: 46.71% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1 (1-3215) x AAR70065 (1-845)

QY	1	CTCCACAACTTCCACCAAGCTCTGCTAGATCCAGGGTGAGGGCGCTATATTTCTGTC	60
DB	306	LeuHisAsnIleProProAsnSerAlaArgSerGlnSerGluGlyProValPheSerCys	325
QY	61	TGCTGGCTCCAGTTCGGGAACAGTAACCTCTGTCCGACTACTGCTCTCCCATATCGTC	120
DB	326	TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal	345
QY	121	AATCTTCTCGAGACTGGGGACCTCGCACCGAACATGGAGAACACAACTCAGGATTCT	180
DB	346	AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	365
QY	181	AGGACCCCTGCTGTTACAGCGGGGTTTTCTCGTTGACAAAGATCCTCACATACC	240
DB	366	ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr	385
QY	241	CGAGACTCTAGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGAGCACCCACGTGTT	299
DB	386	ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlySerThrArgVal-Se	405
QY	300	CTGGCCAAATTCGAGTCCCAACCTCCAACTCCTCACCACTCTGCTCTCCCAATTT	359
DB	405	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuSerSerAsnLe	425
QY	360	GTCTGGCTATCGCTGGATGCTCGGGGTTTTATCATATTCCTTCATCTCTGCTGC	419
DB	425	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaI	445
QY	420	TATGCTCATCTTCTGTTGTTCTTCTGACTTACCAAGTATGTTGCCCGTTGTCCTC	479
DB	445	aMetProHisLeuLeuValGlySerSerGlyLeuSerArgTyrValAlaArgLeuSerSe	465
QY	480	TACTTCCAGGAACATCAACCAACGACGCGGGCCATGCAAGACTGACGACTCTCTGCTC	539
DB	465	rAsnSerArgIleAlaAsnTyrGlnHisGlyThrMetGlnAsnLeuHisAspSerCysSe	485
QY	540	AAGAACTCTAGCTTCTCTCTGTTGCTGTACAAACCTTCGGACGGAACTGCACATT	599
DB	485	rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisIle	505
QY	600	GTATTCCTCATCTCTGCGCTTTCGCAAGATTCCTATGCGAGTGGCGCTCAGTGCC	659
DB	505	uTyrSerHisProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr	525
QY	660	GTTTCTCTGGCTCAGTTTACTAGTCCATTGTTGTCAGTGGTTCGTAGGGCTTTCCCCCA	719
DB	525	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	545
QY	720	CTGTTGGCTTTCAGTTATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	779
DB	545	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisIle	565
QY	780	TGAGTCCCTTTTACCTCTATTACCAATTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTG	839
DB	565	uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	585

QY	840	TAATAAAACCAACCGTTGGGCTACTCCCTTAACTTTCATGCGATATGTAATTTGGAAGTTG	899
DB	585	oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	605
QY	900	GGGTACTTTACCGCAGGAAACATATTGTACTAAACTCAAGCAATTTTTCGAAAACTGCC	959
DB	605	pGlyThrIleProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuPr	625
QY	960	TGTAATAGACTATTGATTGGAAATGTCAAAGAAATTTGGCTCTTTTGGGCTTTC	1019
DB	625	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	645
QY	1020	TGCCCTTTTACACAATGTGGCTATCTCGCTCTGATGCTTATATGATGATATATCAATC	1079
DB	645	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	665
QY	1080	TAAGCAGCTTTTCACTTTCTGCCCAACTTACAAAGCCCTTCTGTGTAAACAATATCTGAA	1139
DB	665	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs	685
QY	1140	CCTTTACCCCTTGGCCGCAACGGTCCGCTCTCTGCCAAGTGTTCGTGACGCAACCCC	1199
DB	685	nLeuTyrProValAlaLahArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr	705
QY	1200	CACTGATGGGCTTGGCCATAGGCCATCAGCGCATGGCTGGAACCTTTCTGGCTCTCT	1259
DB	705	oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe	725
QY	1260	GCCGATCCATCTACGCGAACTCTTAGCAGCTTGTGTTGTCGACGCGTCTGGAGCAAA	1319
DB	725	uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLys	745
QY	1320	ACTTATCGGAACCGCAACTCTGTTGCTCTCTCGGAATAATACACTCTTTCATGGCT	1379
DB	745	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	765
QY	1380	GCTAGGCTGTGTCGCAACTGATCTCTGCGCGGAGCTCTTGTCTAGCTCCCGTCGGC	1439
DB	765	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl	785
QY	1440	GCTGAATCCCGCGGAGCAGCCCGTCTCTGGGGCGCTTTGGGGCTCTACCGTCCCTTCTCA	1499
DB	785	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr	805
QY	1500	TTCTCATCTGCGCGGAGCAGCCCGTCTCTTTCAGCGGCTCTCCCGGTATGTC	1559
DB	805	gLeuProPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	825
QY	1560	TTCTCATCTGCGCGGAGCAGCTCTCTTTCAGCGGCTCTCCCGGTATGTC	1619
DB	825	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	845
QY	1620	G 1620	
DB	845	o 845	
RESULT 8			
ID	ADX40775		
XX	ADX40775 standard; protein; 843 AA.		
AC	ADX40775;		
XX	21-APR-2005 (first entry)		
DT	HBV polymerase protein #18.		
DE	Immune stimulation; polymerase; enzyme.		
XX	Hepatitis B virus.		
OS	WO2005012502-A2.		
XX	10-FEB-2005.		
PD			


```
XX 29-MAR-2004; 2004WO-US009510.
XX PF
XX 28-MAR-2003; 2003US-0458026P.
XX PR
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Cheenut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX DR
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX PS
XX SS Disclosure; Page 380-385; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
XX SQ Sequence 843 AA;

Alignment Scores:
Pred. No.: 1.14e-209 Length: 843
Score: 2750.00 Matches: 517
Percent Similarity: 97.04% Conservative: 8
Best Local Similarity: 95.56% Mismatches: 15
Query Match: 46.55% Indels: 2
DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40775 (1-843)
QY 1 CTCACAACTCCACCAAGCTCTGCTAGATCCCGAGGTGAGGGCCCTATATTTCTCTGC 60
DB LeuHisAsnIleProSerSerAlaArgSerLysSerGluGlyProLeuPheProCys 323
QY 61 TGGTGGCTCCAGTTCGGGAACAGTAACCTGTTCCGACTACTCCCTCTCCCATATCGTC 120
DB TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuThrHisIleVal 343
QY 121 AATCTTCTCGAGGACTGGGACCTCGCACCAATCGGAGAACACACATCAGATTCCT 180
DB AsnLeuLeuGluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro 363
QY 181 AGGACCCCTGCTGCTGTGTACAGGGGGGTTTTCTCGTTGACAGAATCTCTCAATACC 240
DB ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 383
QY 241 GCAGAGTCTAGACTCTG- GTGAGTCTCTCAATTTCTAGGGGAGACCCACGTTTC 299
DB ThrGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnThrGlnVal-Se 403
QY 300 CTGGCCAAAATTCGACGTCCCACTCCATCACTCACCACTCTGTGTCCTCAATTT 359
DB rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 423
QY 360 GTCTGTGCTATCGCTGGATGTGTCTGGGGGTTTTATCATATTCCTCTTCATCTGCTGC 419
DB uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl 443
QY 420 TAGCCCTCATCTTCTGTTGGTCTTCTGGACTACCAAGGTATGTCCTGTCCTC 479
DB aMetProHisLeuLeuValGlySerSerGlyLeuGlnArgTyrValAlaArgLeuSerSe 463
QY 480 TACTTCCAGGAACATCAACACGACGCGGCCATCGACAGCTGACGACTCTCTGCTC 539
DB
```

```
Db 463 rThrSerArgAsnIleAsnTyrGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe 483
QY 540 AAGGAAACTCTACGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 599
DB 543 rLysHisLeuTyrValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 503
QY 600 GTATTCCCATCCCATCATCTCTGGGCTTTCGCAAGATCTCTATGGAGTGGGCCCTCAGTCC 659
DB 503 uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 523
QY 660 GTTCTCTCTGCTCAGTTTACTAGTCGCAATTTGTTCTAGTGGTTCGTAGGGCTTTCCCCCA 719
DB 523 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValCysArgAlaPheProHi 543
QY 720 CTGTTTCTGCTTTCAGTTATATGATGTGGTATTGGGGCGGAAGTCTGTACCAACATCT 779
DB 543 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisIle 563
QY 780 TGAGTCCCTTTTACCTCTATTACCAATTTTCTTTTCTTTTGGGTATACATTTTAAACCC 839
DB 563 uGluSerLeuPheThrSerIleThrAsnPheMetLeuSerLeuGlyIleHisLeuAsnPr 583
QY 840 TAATAAAACCAACGTTGGGGCTACTCCCTTAACCTTATCTGCGATATGTAATGGAAGTTG 899
DB 583 oAsnLysThrLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr 603
QY 900 GGGTACTTTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACATGCC 959
DB 603 pGlyThrLeuProGlnGluHisIleValGlnLysLeuLysGlnCysPheArgLysLeuPr 623
QY 960 TGTAAATPAGACCTTATGATTGGAAAGTATGTCAAGAAATGTGGGTCTTTTGGGCTTTC 1019
DB 623 oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 643
QY 1020 TGCCCTTTTACCAATGTGGCTATCTGCTTGTATGCTTATATGATGATATACATC 1079
DB 643 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 663
QY 1080 TAAGCAGGCTTTCTACTTTCTCGCAACTTACAAAGCCTTTCTGTGTAAACAATATCTGAA 1139
DB 663 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 683
QY 1140 CTTTATCCCGCTGCTCCCGCAACGGTCCGCTCTCTGCAAGTGTTCGTACGCAACCC 1199
DB 683 nLeuTyrProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrPr 703
QY 1200 CACTGGATGGGGCTTGGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCC 1259
DB 703 oThrGlyTrpGlyLeuAlaIleGlyHisGlnArgMetArgGlyThrPheValAlaProLe 723
QY 1260 GCCGATCCATACGCGGAACCTCTAGCAGCTTGTGTTTGTGCGACGCGGTCTGGAGCAA 1319
DB 723 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaIle 743
QY 1320 ACTTATCGGAACCGCAACACTCTGTGTCTCTCTCGGAAATACACCTCTTCCATGGCT 1379
DB 743 rLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe 763
QY 1380 GCTAGGTGTGCTGCCAATCGGATCTCGCGGGAGCGTCTTGTCTACGTCCCGTGGC 1439
DB 763 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 783
QY 1440 GCTGAATCCCGCGCAGCACCCGCTCTCGGGCGGCTTGTGGGCTCTTACCGTCCCTCTTCA 1499
DB 783 aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuAr 803
QY 1500 TCTGCCGCTTCCGCGCAGCACCGGGCGCACCTCTCTTTACGCGGTCTCCCGTATGTGCC 1559
DB 803 gLeuSerPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr 823
QY 1560 TTCTCATCTGCGGACCGTGTGCACTTCGCTTACCTCTCTGACGTCTGCGATGGAGCACCC 1619
DB 823 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr 843
```


QY	1620 G 1620	61	TGTTGGCTCCAGTTCCGGAACAGTAAACCCCTGTTCCGACTACTGCTCTCCATATCGTC	120
DB	843 O 843	21	TrpTrpLeuGlnPheArgAsnSerIysProCysSerAspTyrCysLeuThrHisIleVal	40
QY	ADME68212	121	AATCTTCTCGAGGACTGGGACCCCTGACACCAACATGAGAGAACCAACATCAGGATTCCT	180
DB	ADME68212	41	AsnLeuLeuGluAspTyrGlyProCysThrGluHisGlyGluHisAsnIleArgIlePro	60
QY	03-JUN-2004 (first entry)	181	AGGACCCCTGCTCGTTTACAGGGGGGTTTCTCGTTGACAAAGATCCCTCACAAATACC	240
DB	GRP94 protein related Hepatitis B virus protein, SEQ ID No 4.	61	ArgThrProAlaArgValThrGlyValPheLeuValAspLysAsnProHisAsnThr	80
QY	GRP94; chaperone; DNA polymerase; Hepatitis B virus; HBV.	241	GCAGAGTCTAGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGACACCCACCTGTTTC	299
DB	Hepatitis B virus.	81	ThrGluSerArgLeuValAlaAspPheSerGlnPheSerArgGlySerThrHisVal-Se	100
QY	KR2002061721-A.	300	CTGGCCAAAATTGCGAGTCCCAACCTCCAATCACTCAACCACTCTTGTCTCTCAATTT	359
DB	25-JUL-2002.	100	rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe	120
QY	17-JAN-2001; 2001KR-00002647.	360	GTCTGGCTATCGCTGGATGTCTCGGGCGTTTATCATATTTCTCTCTCTCTCTCTCTCC	419
DB	17-JAN-2001; 2001KR-00002647.	120	uSerTrpLeuSerLeuAspValSerAlaAlaPheTyrHisIleProLeuHisProAlaAl	140
QY	(MEDE-) MEDEXBIO CO.	420	TATGCTCATCTTCTTGTGTTCTCTGGACTACCAAGTATGTTCGCCGTTTGTCTCTC	479
DB	Cho YH, Kim SS, Noh HM, Shin HJ;	140	amMetProHisLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSe	160
QY	WPI; 2003-117827/11.	480	TACTTCCAGGAACATCAACACACGACGCGGCCCATGCAAGACCTGCACGACTCTCTGC	539
DB	N-PSDB; ADM68211.	160	rThrSerArgAsnIleAsnHisGlnHisGlyThrMetGlnAspLeuHisAspSerCysSe	180
QY	Grp94 acting as chaperone of DNA polymerase of hepatitis b virus and	540	AAGGAAACTTACGTTTCCCTCTTCTGCTGTACAAAACCTTCGGACGGAACCTGCACAT	599
DB	expression vector system for preparing a DNA polymerase in active state	180	rArgAsnLeuTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLe	200
QY	of hepatitis b virus.	600	GTATTCCCATCCCATCATCTCGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC	659
DB	Disclosure; SEQ ID NO 4; 15pp; Korean.	200	uTyrSerHisProIleIleLeuGlyPheArgLysIleProMetGlyGlyGlyLeuSerPr	220
QY	The invention relates to a novel GRP94 protein. The GRP94 protein acts as	660	GTCTTCTCTGGCTCAGTTTACTAGTCCCATTTGTTTCAGTGGTTCGTAGGGCTTCCCCCA	719
DB	a chaperone of DNA polymerase of hepatitis B virus. The invention further	220	oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHi	240
QY	relates to: an expression vector system for preparing a DNA polymerase in	720	CTGTTTGGCTTTCAGTTATATGATGATGATGATGATGATGATGATGATGATGATGATG	779
DB	the active state of a hepatitis B virus; a transformant and a	240	sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe	260
QY	manufacturing method of DNA polymerase of active hepatitis B virus using	780	TCAGTCCCTTTTACTCTATTACCAATTTCTTTTGTCTTTTGGGTATACATTTAAACCC	839
DB	the transformant, and rapidly screening potential matter for the	260	uGluSerLeuPheThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr	280
QY	prevention and treatment of diseases associated with hepatitis B virus.	840	TAATAAACCACAAACGTTGGGCTACTCCCTTAACTTTCATGGGATATGTAATTTGGAAGTTG	899
DB	The manufacturing method of DNA polymerase of active hepatitis B virus	280	oAsnLysThrLysArgTyrGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTr	300
QY	comprises transforming a host cell simultaneously with an expression	900	GGGTACTTTTACCGCAGGAACATATTGTACTATAAACCTCAAGCAATGTTTTCGAAAACTGCC	959
DB	vector including a coding gene of hepatitis B virus DNA polymerase and an	300	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	320
QY	expression vector under the condition of GRP94 protein, culturing the	960	TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGTGGCTTTTGGCTTTC	1019
DB	transformed host under the condition to initiate replication at two of	320	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	340
QY	replication origins of the expression vectors, and purifying DNA	1020	TGCCCTTTTACACAAATGTGGCTATCTCTGCTTGTATGCTCTTATATGATGATATACAAATC	1079
DB	polymerase of hepatitis B virus included in the culture solution. This	340	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	360
QY	sequence represents a Hepatitis B virus protein relating to the GRP94	1080	TAAGCAGGCTTTCATCTTCTCGCCAACTTCAAGGCTTCTCTGTGTAAACAATATCTGAA	1139
DB	protein of the invention.	360	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi	380
QY	Seq	1140	CCTTTACCCGTTGCGCGGCAACGGTCCGGTCTCTGCCAAGTCTTTGCTGACGCAACCCC	1199

Alignment Scores:
 Pred. No.: 9.9e-209 Length: 539
 Score: 2737.50 Matches: 517
 Percent Similarity: 96.86% Conservative: 7
 Best Local Similarity: 95.56% Mismatches: 15
 Query Match: 46.34% Indels: 3
 DB: 7 Gaps: 1

US-10-761-006A-1 (1-3215) x ADM68212 (1-539)
 QY 1 CTCACAACTTCCACCAAGCTCTGCTAGATCCAGGCTGAGGGGCTATATTTCTGTC 60
 DB 1 LeuHisIleSerProSerProAlaArgSerGlnSerGlnGlyProIlePheSerSer 20

QY	840	TAATAAAACCAAACGTTGGGGCTACTCCCTTAACTTCATGGGATATGTATATGGAAGTTG	899
DB	583	oAnLysThrLysArgTgPglyTyrSerLeuAenPheMetGlyTyrValIleGlySerTr	603
QY	900	GGGTACTTTACCCGAGAACATATTGCTATAAACTCAAGCATGTCTTCGAAAACGTGCC	959
DB	603	pGlyThrLeuProGlnGluHisIleValLeuLysIleLysGlnCysPheArgLysLeuPr	623
QY	960	TGTAATAATAGACCTATTGATTGGAAAGTATGCTCAAGAATTTGGTGGCTTTTGGCTTTGC	1019
DB	623	oValAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl	643
QY	1020	TGCCCTTTTACAAATGTGGCTATCTGCTGTGATGCTTCCTTATATGCGATGATATACAATC	1079
DB	643	aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe	663
QY	1080	TAGCAGGCTTTCACCTTCTCGCCCACTTACAGGCTTCTCTGTGTAAACAATATCTGAA	1139
DB	663	rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHi	683
QY	1140	CTTTACCCCGTTGGCCGGCAACGGTCCGGCTCTCTGCCAAGTGTTCGTGACCAACCCC	1199
DB	683	sLeuTyrProValAlaArg--ArgThrAlaLeuCysGlnValPheAlaAspAlaThrPr	702
QY	1200	CACGTGGATGGGCTTGGCCATAGCCCATCAGCCGATGGCTGGAACTTCCTTCGTCTCTCT	1259
DB	702	oThrGlyTgPglyLeuAlaIleGlyHisArgMetArgGlyThrPheValAlaProLe	722
QY	1260	GCGCATCCATACTCGCGGAATCTCTAGCAGCTGTGTTCGTGCGACGGCTCTGGAGCAA	1319
DB	722	uProIleHisThrAlaGluLeuLeuAlaIaCysPheAlaArgSerArgSerGlyAlaLy	742
QY	1320	ACTTATCGGAACCGACAACCTGTGTGTCTCTCTCGGAAATACACCTTCCTTCATGGCT	1379
DB	742	sLeuIleGlyThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLe	762
QY	1380	GCTAGGTGTGTCGCCAACTCGATTCCTGCGCGGACGTCTCTTGTCTACGTCGGTGGCC	1439
DB	762	uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrTyrPheValTyrValProSerAl	782
QY	1440	GCTGAATCCCGCGAGACACCGCTCTCGGGCCGTTTGGGGCTCTACCGTCCCTTCTTCA	1499
DB	782	aLeuAsnProAlaAspAspProSerArgGlyArgLeuGlyLeuLeuArgProLeuLeuHi	802
QY	1500	TCTGCCGTTCGGCCGACCGGGCGCACCTCTCTTTTACGCGGTCTCCCGCATGTGCC	1559
DB	802	sLeuArgPheArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValPr	822
QY	1560	TTCTCATCTGCGGACCGTGTGCATCTCGCTTCTCACTCTGCACGTCTGCATGGAGACCAC	1619
DB	822	oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPr	842
QY	1620	G 1620	
DB	842	o 842	

RESULT 11

ADX40774
ID ADX40774 standard; protein; 825 AA.

AC ADX40774;

XX

DT 21-APR-2005 (first entry)

XX

DE HBV polymerase protein #17.

Immune stimulation

XX
MY
XXXXX

OS Hepatitis B virus

3 XX

PN WO2005012502-A2.

XX

PD 10-FEB-2005.
 XX
 XX 29-MAR-2004; 2004WO-US009510.
 PF
 XX
 XX 28-MAR-2003; 2003US-0458026P.
 PR
 XX (EPIM-) EPIMUNE INC.
 XX
 PA Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 PI WPI; 2005-132661/14.
 XX
 XX
 DR Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX
 XX Disclosure; Page 380-385; 458pp; English.
 PS
 XX The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HBV polymerase protein used in the scope of the invention.
 XX
 XX
 SQ Sequence 825 AA;

Alignment Scores:		
Pred. No.:	9.76e-205	825
Score:	2688.00	504
Percent Similarity:	95.92%	13
Best Local Similarity:	93.51%	21
Query Match:	45.51%	0
DB:	9	2
DB:		

US-10-761-006A-1 (1-13215) x ADX40774 (1-825)

Db 447 rargasnleasn***GlnHISgLyThrMetGlnAspLeuHisaspSerCysSerLysHi 467
Qy 546 ACTTACGTTTCCCTCTTGTGCTGTACAAACCTTCGAGCGGAACTGCACCTGTATTTC 605
Db 467 sLeuTyRValSerLeuLeuLeuTyRLeuPheGlyArgLysLeuHisLeuTyRSe 487
Qy 606 CCATCCCATCATCTCGGGCTTTCGCAAGATTCCTATGGAGTGGGCCCTCAGTCGCTTCT 665
Db 487 rHisProIle***LeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLe 507
Qy 666 CCTGCTCAGTTACTAGTGCATTTGTTCTAGTGGTGTGGCTTCCGCCACTGTTT 725
Db 507 uLeuAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLe 527
Qy 726 GGCTTTTCAGTTATATGATGATGTGTTATGGGGCGAAGTCTGTACAACTCTTGTAGTC 785
Db 527 uAlaPheSerTyRMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluAl 547
Qy 786 CTTTTTTACCTCTATTACCAATTTCTTTTCTTTGGGTATATACATTTAAACCTTAATAA 845
Db 547 aLeuTyRThrSerIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProAsnLy 567
Qy 846 AACCAACGTTGGGCTACTCCCTTAACCTTCATGGGATATGATTTGAAGTTGGGTAC 905
Db 567 sThrLysArgTrpGlyTyRSerLeuAsnPheMetGlyTyRValIleGlySerTrpGlyTh 587
Qy 906 TTTACCGCAGGAACATATTGTACTAAACTCAAGCAATGTTTTCGAAACTGCCTGTAAA 965
Db 587 rLeuProGlnAspHisIleValGlnLysLeuLysGlnCysPheArgLysLeuProValAs 607
Qy 966 TAGACCTATTGATGGAAAGTATGTCAAGAAATTTGTGGTCTTTTGGCTTTGTCGCCCC 1025
Db 607 nArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAlaAlaPr 627
Qy 1026 TTTTACCAAGTGGCTATCTGCTTGCCTTGCATGCTTATATGATCATGATATACATCTAAGCA 1085
Db 627 oPheThrGlnCysGlyTyRProAlaLeuMetProLeuTyRAlaCysIleGlnSerLysGl 647
Qy 1086 GGCTTTTCACTTTCTCGCAACTTACAGGCTTCTGTGTAAACAATATCTGAACCTTTA 1145
Db 647 nAlaPheThrPheSerProThrTyRLeuAlaPheLeuCysLysGlnTyRLeuAsnLeuTy 667
Qy 1146 CCCGTTGCCCGCAACGGTCCGGTCTCTGCAAGTGTTCGCAAGTGTTCGCAACCCCACTGG 1205
Db 667 rProValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGl 687
Qy 1206 ATGGGGCTTGGCATAGCCATCAGCGCATCGCTGGAACCTTCTGGCTCCTCTCCGAT 1265
Db 687 yTrpGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIl 707
Qy 1266 CCATACCTCGGAACTCCTAGCAGCTTGTGTCGAGCGGCTGAGCAAACTTAT 1325
Db 707 eHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIl 727
Qy 1326 CGGAACCGCAACTCTGTGTCTCTCTCGGAATATACACCTCTTCCATGGCTGCTAGG 1385
Db 727 eGlyThrAspAsnSerValValLeuSerArgLysTyRThrSerPheProTrpLeuLeuGl 747
Qy 1386 GTGTGCTGCCAACTGGATCCTCGCGGAGCTCTTGTGTCTAGCTCCGCTCGGCGCTGAA 1445
Db 747 yCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyRValProSerAlaLeuAs 767
Qy 1446 TCCGCGGACGACCGCTCTCGGGGCGGTTTGGGGCTCTACCGTCCCTCTTCTCATCTGCC 1505
Db 767 nProAlaAspAspProSerArgGlyArgLeuGlyLeuTyRArgProLeuLeuHisLeuPr 787
Qy 1506 GTTCGGCGGACCAACGGGCGCACTCTCTTTTACCGGCTCTCCCGTATGTCCTTCTCA 1565
Db 787 oPheArgProThrThrGlyArgThrSerLeuTyRAlaValSerProSerValProSerHi 807
Qy 1566 TCTCGGCGGCTGTGCACTTGGCTTCACTCTGCACTGCGTCCGATGAGACCAACG 1620

Db 807 sLeuProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 825
RESULT 12
ADX40773
ID ADX40773 standard; protein; 827 AA.
AC ADX40773;
XX 21-APR-2005 (first entry)
XX HBV polymerase protein #16.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis B virus.
XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 380-385; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HBV polymerase protein used in the scope of the invention.
SQ Sequence 827 AA;
Alignment Scores:
Pred. No.: 5.57e-204 Length: 827
Score: 2678.50 Matches: 504
Percent Similarity: 95.91% Conservative: 12
Best Local Similarity: 93.68% Mismatches: 18
Query Match: 45.34% Indels: 5
DB: Gaps: 1
US-10-761-006A-1 (1-3215) x ADX40773 (1-827)
Qy 19 AGCTCTCTAGATCCAGGGTGAGGGG-----CCTATATTTTCTGCTGTGGCTC 69
Db 291 ThrSerLysArgGlnSerSerSerGlyHisAlaValGluLeuHisProCysTrpTrpLeu 310
Qy 70 CAGTTCCGGAAACAGTAAACCTGTTCCGACTACTGCTCTCCCATATCGTCAATCTTCTC 129
Db 311 GlnPheArgAsnSerLysProCysSerAspTyRcysLeuThrHisIleValAsnLeuLeu 330
Qy 130 GAGGACTGGGACCTCCACCGCAACATGGAGAACACAACTCAGGATTCCTAGACCCCT 189
Db 331 GluAspTrpGlyProCysThrGluHisGlyGluHisAsnIleArgIleProArgTrpPro 350
Qy 190 GCTCGTCTTACAGCGGGGTTTCTGTTGACAGAGATCTCTCAATACCGCAGAGTCT 249
Db 351 AlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThrThrGluSer 370

QY 250 AGACTCTG-GTGGACTTCTCTCAATTTCTAGGGGAGCACCACGCTGTTCTCTGGCCAAA 308
Db |||||
QY 371 ArgLeuValValAspPheSerGlnPheSerArgGlySerThr***Val-SerTrpProLy 390
Db |||||
QY 309 ATTGCGAGTCCCAACCTCCAACTACTCACCACCTCTTGTCTCCATTTGTCTCTGGCT 368
Db |||||
QY 330 sPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLeuSerTrpLe 410
Db |||||
QY 369 ATCGCTGGATGTCTGGGGCGTTTATCATATATCTCTTTCATCTGCTGCTATGCCCTCA 428
Db |||||
QY 410 uSerLeuAspValSerAlaAlaPheTyRHisileProLeuHisProAlaAlaMetProHi 430
Db |||||
QY 429 TCTTCTGTTGGTCTTCTGAGTACCAAGGTATGTTGCCGTTTGTCTCTACTTCCAG 488
Db |||||
QY 430 sLeuLeuValGlySerSerGlyLeuProArgTyrValAlaArgLeuSerSerThrSerAr 450
Db |||||
QY 489 GAACATCAACACACGAGCGGGCCATGCAAGACCTGCAGCACTCTCTCAAGGAAACT 548
Db |||||
QY 450 GABN***AsnTyrGlnHisGlyThrMetGlnAspLeuHisGlnSerCysSerArgAsnLe 470
Db |||||
QY 549 CTACGTTTCCCTCTTGTCTGTACAAAACCTTCGAGCGAACTGCACCTTGTATTTCCCA 608
Db |||||
QY 470 uTyrValSerLeuLeuLeuTyrLysThrPheGlyArgLysLeuHisLeuTyrSerHi 490
Db |||||
QY 609 TCCCATCATCTGGGCTTTCGCAAGATTCCTATGGAGTGGCCCTCAGTCGTTCTCTCT 668
Db |||||
QY 490 sProIleLeuGlyPheArgLysIleProMetGlyValGlyLeuSerProPheLeuLe 510
Db |||||
QY 669 GCGTCAGTTACTAGTGCATTTGTTTCAGTGTCTGTAGGGCTTCCCGCTGTTTGGC 728
Db |||||
QY 510 uAlaGlnPheThrSerAlaIleCysSerValValArgArgAlaPheProHisCysLeuAl 530
Db |||||
QY 729 TTTTCAGTTATATGATGTGTGTATTTGGGGCGAAGTCTGTACAACTCTTGAGTCCT 788
Db |||||
QY 530 aPheSerTyrMetAspAspValValLeuGlyAlaLysSerValGlnHisLeuGluSerLe 550
Db |||||
QY 789 TTTTACCTCTATACAAATTTCTTTTGTCTTTGGGTATACATTTAAACCTTAATAAAC 848
Db |||||
QY 550 uPheThrAlaIleThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnProHisLysTh 570
Db |||||
QY 849 CAACGTTGGGCTACTCCCTAACTTCACTGAGATATGTAATTTGGAAGTTGGGGTACTTT 908
Db |||||
QY 570 rLysArgTrpGlyTyrSerLeuAsnPheMetGlyTyrValIleGlySerTrpGlyThrLe 590
Db |||||
QY 909 ACCGAGGAAACATATTGTACTAAACTCAAGCAATGTTTTCGAAAACCTGCTGTAATAG 968
Db |||||
QY 590 uProGlnGluHisIleValGlnLysIleLysGlnCysPheArgLysLeuProValAsnAr 610
Db |||||
QY 969 ACTATTTGATGGAAAGTATGCAAGAAATTTGTGGGTCTTTTGGGCTTGTCTGCCCTTT 1028
Db |||||
QY 610 gProIleAspTrpLysValCysGln***IleValGlyLeuLeuGlyPheAlaAlaProPh 630
Db |||||
QY 1029 TACCAATGTGCTATCTCGCTCTGATGCTTTATATGATCATGTATACATCTCAAGCAGC 1088
Db |||||
QY 630 eThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnAlaLysGlnAl 650
Db |||||
QY 1089 TTTTCACTTTCTCGCCAACTTACAAGGCTTTCTGTGTAAACAATATCTGAACCTTTTACC 1148
Db |||||
QY 650 aPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuHisLeuTyrPr 670
Db |||||
QY 1149 CTTTGGCCCGGCAAGCTCGGTCTCTGCAAGTGTGTCGACGCAACCCCACTGGAG 1208
Db |||||
QY 670 oValAlaArgGlnArgSerGlyLeuCysGlnValPheAlaAspAlaThrProThrGlyTr 690
Db |||||
QY 1209 GGGCTGGCCATAGGCGATCAGCGATGCTGGAACCTTCTGGCTCTCTGCCATCCA 1268
Db |||||
QY 690 pGlyLeuAlaIleGlyHisArgArgMetArgGlyThrPheValAlaProLeuProIleHi 710
Db |||||
QY 1269 TACTGGGAACTCTTACGAGCTTGTTTGCTGGCAGCGCTTGGAGCAAACTTATCCG 1328
Db |||||
QY 710 sThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaLysLeuIleGl 730
Db |||||

QY 1329 AACGACAACTCTGTTGTCTCTCTCTCGGAAATACACCTCTTTTCATGGCTGCTAGGTG 1388
Db |||||
QY 730 yThrAspAsnSerValValLeuSerArgLysTyrThrSerPheProTrpLeuLeuGlyCy 750
Db |||||
QY 1389 TGCTGCCAACTGGATCTCTGGCGGAGCTCTTTTGTCTACGTCCGTCGGCGCTGAATCC 1448
Db |||||
QY 750 sAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAlaLeuAsnPr 770
Db |||||
QY 1449 CGCGAGCAGCCGCTCTGGGGCCGTTTGGGCTCTACGCTCCCTTCTTCATCTGCCGTT 1508
Db |||||
QY 770 oAlaAspAspProSerArgGlyArgLeuGlyLeuTyrArgProLeuLeuHisLeuProPh 790
Db |||||
QY 1509 CCGCGCCGACACGCGGCGCACCTCTCTTACGCGCTCTCCCGTATGCTATGCTTCTCATCT 1568
Db |||||
QY 790 eArgProThrThrGlyArgThrSerLeuTyrAlaValSerProSerValProSerHisLe 810
Db |||||
QY 1569 GCCCGACGCTGTGCATCTTCGCTTTCACCTCTGCACGTCTGCATGGAGACACCG 1620
Db |||||
QY 810 uProAspArgValHisPheAlaSerProLeuHisValAlaTrpArgProPro 827
Db |||||
RESULT 13
AAE04708
ID AAE04708 standard; protein; 842 AA.
XX
AC AAE04708;
XX
DT 04-SEP-2001 (first entry)
XX
DE Hepatitis B virus FRI strain genotype G HBpol protein.
XX
KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
KW preS2; surface antigen; HBeAg; HBx protein; vaccine; liver disease;
KW hepatitis; liver cancer; HBeAg; core antigen.
XX
OS Hepatitis B virus.
XX
PN WO200138498-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032108.
XX
PR 24-NOV-1999; 99US-0167206P.
XX
PA (PHAR-) PHARMASSET INC.
XX (INNO-) INNOGENETICS NV.
PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
PI Rousseau R;
XX
DR WPI; 2001-367676/38.
DR N-PSDB; AAD09091.
XX
PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
XX Example; Fig 6; 84pp; English.
PS
CC The present invention relates to hepatitis B virus (HBV) strain FRI,
CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC PreS2 and surface antigen HBeAg) and HBx proteins. HBV genotype G nucleic
CC acids and polypeptides are useful for diagnosing, prognosing and treating
CC infections caused by HBV genotype G. They can be used in a vaccine to
CC treat or prevent HBV genotype G infection. The HBV genotype G derived
CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC sample or diagnosis of HBV genotype G infection. The presence of HBV
CC genotype G statistically correlates with the presence of liver damage
CC and/or liver cancer in the subject. The HBV genotype G core insert
CC peptide encoding nucleic acid is useful for designing monitoring assays
CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC antibodies or antigens of HBV genotype G are useful for identifying a


```
QY 1500 TCTGCGCTTCGGCCGACACCGGGCGCACCCTCTCTTTACGGGTCTCCCGTATGTC 1559
Db 802 gLeuProPheLeuProThrThrGlyArgThrSerLeuTyzAlaValSerProSerValPr 822
QY 1560 TTCTCATCTCCGACCGCTGCACTTCGCTTCACCTCTGCACTGCGCATGGAGACCACC 1619
Db 822 oSerHisLeuProAspArgValHisPheAlaSerProLeuHisValThrTrpLysProPr 842
QY 1620 G 1620
Db 842 o 842

RESULT 15
ADX40761
ID ADX40761 standard; protein; 832 AA.
XX
AC ADX40761;
XX
DT 21-APR-2005 (first entry)
XX
DE HBV polymerase protein #4.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis B virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 380-385; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HBV polymerase protein used in the scope of the invention.
XX
SQ Sequence 832 AA;

Alignment Scores:
Pred. No.: 4.97e-202 Length: 832
Score: 2654.00 Matches: 492
Percent Similarity: 95.56% Conservative: 25
Best Local Similarity: 90.94% Mismatches: 23
Query Match: 44.93% Indels: 2
DB: 9 Gaps: 0

US-10-761-006A-1 (1-3215) x ADX40761 (1-832)

QY 1 CTCCACCAACATTCACCAAGCTCTGCTAGATCCAGGCTGAGGGCCCTATATTTCTCGC 60
Db 293 LeuHisAnPheProProAnSerAlaArgSerGlnGlyGluArgProValPheProCys 312
QY 61 TGGTGGCTCCAGTTCGCGAACAGTAACCCCTGTTCCGACTACTCCCTCTCCCATATCGTC 120
```

```
313 TrpTrpLeuGlnPheArgAsnSerLysProCysSerAspTyrCysLeuSerHisLeVal 332
QY 121 AATCTTTCTCGAGACATGGGACCCCTCACCGAACATGGAGAACACACACATCAGATTTCCT 180
Db 333 AsnLeuLeuGlnAspTrpGlyProCysThrGluHisGlyGluHisLeArgGlePro 352
QY 181 AGGACCCCTGCTCGTGTATACAGCGGGGTTTTCTCGTTGACAAAGATCCTCAAAATACC 240
Db 353 ArgThrProAlaArgValThrGlyGlyValPheLeuValAspLysAsnProHisAsnThr 372
QY 241 CGAGAGCTAGACTCTG-GTGACTTCTCTCAATTTTCTAGGGGAGACCCACAGTGTTC 299
Db 373 AlaGluSerArgLeuValValAspPheSerGlnPheSerArgGlyAsnHisArgVal-Se 392
QY 300 CTGGCCAAAATTCGACAGTCCCCAACCTCCAAATCACCTCACCAACCTCTGTCTCCCAATTT 359
Db 392 rTrpProLysPheAlaValProAsnLeuGlnSerLeuThrAsnLeuLeuSerSerAsnLe 412
QY 360 GTCTTCGGCTATCGCTGGATGTGTCTGGCGGTTTTATCATATTTCTCTTCTCATCTGCTGC 419
Db 412 uSerTrpLeuSerLeuAspValSerAlaPheTyrHisLeuProLeuHisProAlaSe 432
QY 420 TATGCTCATCTCTCTGTTGGTCTTCTTGACTACCAAGTATGTTGCCGCTTGTCTCCTC 479
Db 432 rMetProHisLeuLeuValGlySerThrGlyLeuSerArgTyrValAlaArgValSerSe 452
QY 480 TACTTCCAGGAACATCAACACACGACGCGGGCCATGCAAGACCTGCACCACTCTGCTC 539
Db 452 rAsnSerArgLlePheAsnHisGlnArgGlyThrMetGlnAsnLeuHisAspTyrCysSe 472
QY 540 AAGGAAACTCTACGTTTCCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db 472 rArgAsnLeuTyrValSerLeuLeuLeuLeuTyrGlnThrPheGlyArgLysLeuHisLe 492
QY 600 GTATTCCTCCATCCCATCATCTGGGCTTTCGCAAGATTCCTATGGGAGTGGGCTCAGTCC 659
Db 492 uTyrSerHisProIleLeuLeuGlyPheArgLysIleProMetGlyValGlyLeuSerPr 512
QY 660 GTTTCCTCGGCTCAGTTTACTAGTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Db 512 oPheLeuLeuAlaGlnPheThrSerAlaIleCysSerValValArgAlaPheProHi 532
QY 720 CTGTTTGGCTTTCAGTTTATATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
Db 532 sCysLeuAlaPheSerTyrMetAspValValLeuGlyAlaLysSerValGlnHisLe 552
QY 780 TGAGTCCCTTTTACCTCTATATACCAATTTTCTTTTGTCTTTGGGTATATACATTTAAACCC 839
Db 552 uGluSerLeuPheThrAlaValThrAsnPheLeuLeuSerLeuGlyIleHisLeuAsnPr 572
QY 840 TAATAAAACCAACGTTGGGGCTACCTCCCTTAACTTTCATGGGATATGTAATTTGGAAGTTG 899
Db 572 oAsnLysThrLysArgTrpGlyTyrSerLeuHisPheMetGlyTyrValIleGlySerTy 592
QY 900 GGGTACTTTTACCCGAGAACATATTTGTACTTAAACCTCAAGCAATGTTTTCGAAACATGCC 959
Db 592 rGlySerLeuProGlnAspHisIleIleGlnLysIleLysGluCysPheArgLysLeuPr 612
QY 960 TGTAAATAGACCTATTGATTGGAAAGTATGTCAAGAAATTTGGGTCTTTTGGGCTTTGTC 1019
Db 612 oIleAsnArgProIleAspTrpLysValCysGlnArgIleValGlyLeuLeuGlyPheAl 632
QY 1020 TGCCCCCTTTTACACAAATGTGGCTATCTGCTGCTGTGATGCCCTTTATATCATGTATACAAATC 1079
Db 632 aAlaProPheThrGlnCysGlyTyrProAlaLeuMetProLeuTyrAlaCysIleGlnSe 652
QY 1080 TAAAGCAGGCTTTTCACTTTTCTCGCCAACTTAAAGGCGCTTCTTGTGTAAACAATATCTGAA 1139
Db 652 rLysGlnAlaPheThrPheSerProThrTyrLysAlaPheLeuCysLysGlnTyrLeuAs 672
QY 1140 CTTTACCCGCTTTCGCCCGCAACGGTCCGCTCTGCGCAAGTGTTCGTACGCAACCCC 1199
```



```

Db      672 nLeuTyPrValAlaArgGlnArgProGlyLeuCysGlnValPheAlaAspAlaThrPr 692
QY      1200 CACTGGATGGGCTTGGCCATAGGCCATAGCGCATGGCTGGAAACCTTTCTGGCTCTCT 1259
Db      692 oThrGlyTrpGlyLeuValMetGlyHisGlnArgMetArgGlyThrPheLeuAlaProLe 712
QY      1260 GCCGATCCATCTACTGGGAACTCCTAGCAGCTTGTCTGCTCGCAGCCGGTCTGGAGCAA 1319
Db      712 uProIleHisThrAlaGluLeuLeuAlaAlaCysPheAlaArgSerArgSerGlyAlaAs 732
QY      1320 ACTTATCGGAACCGCAACTCTGTTGTCTCTCTCGGAAATACACCTCTTTCCATGGCT 1379
Db      732 nIleLeuGlyThrAspAsnSerValIleLeuSerArgIysTyThrSerPheProTrpLe 752
QY      1380 GCTAGGATGCTGCAACTGGATCCTCGCGGACGTCCTTTGTCTACGTCCCGTCGGC 1439
Db      752 uLeuGlyCysAlaAlaAsnTrpIleLeuArgGlyThrSerPheValTyrValProSerAl 772
QY      1440 GCTGAATCCCGCGGACGCGCTCTCGGGCGGTTTGGGGCTCTACCGTCCGCTCTTCA 1499
Db      772 aLeuAsnProAlaAspAspProSerArgGlyIlePheArgProLeuLeuAr 792
QY      1500 TCTGCCGTTCCGGCGGACCGCGGCGCACCTCTCTTTACGGGTCTCCCGTATGTGCC 1559
Db      792 gLeuProPheArgProThrThrGlyArgThrSerLeuTyAlaAspSerProSerValPr 812
QY      1560 TTCTCATCTCGCGGACCGGTGTGCACTTCGCTTTCACCTCTGCACGTCGCATGGAGACCAC 1619
Db      812 oSerHisLeuProValArgValHisPheAlaSerProLeuHisValAlaTTrpArgProPr 832
QY      1620 G 1620
Db      832 o 832

```

Search completed: December 1, 2005, 23:35:15
Job time : 520.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 27, 2005, 20:42:07 ; Search time 3.91 Seconds
(without alignments)
3395.886 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595
Perfect score: 139
Sequence: 1 ACACCTCTGCTCAGGAAA.....AACCTTCGGACGGAACCTGC 69

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgm2_1/USPTO_epool_p/US10761006/runat_27122005_192901_15763/app_query.fasta_1.1102
-DB_PIR -QFMT-fastan -SUFFIX=rpr -MINMATCH=0.1 -LOQPCI=0 -LOQEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRF=US10761006 @CGN 1 1 77 @runat_27122005_192901_15763 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133	95.7	226	2 JQ2054	surface antigen -
2	130	93.5	226	1 JQ1577	major surface anti
3	130	93.5	226	1 SAVLHV	major surface anti
4	130	93.5	226	2 JQ2052	surface antigen -
5	130	93.5	226	2 JQ2053	surface antigen -
6	130	93.5	226	2 JQ2050	surface antigen -
7	130	93.5	226	2 JQ2048	surface antigen -
8	130	93.5	226	2 JQ2055	surface antigen -
9	130	93.5	226	2 JQ2078	surface antigen -
10	130	93.5	226	2 JQ2046	surface antigen -
11	130	93.5	226	2 JQ2045	surface antigen -
12	130	93.5	226	2 JQ2056	surface antigen -
13	130	93.5	389	1 SAVLVE	large surface anti
14	130	93.5	400	1 JQ1575	major surface anti

15	130	93.5	400	1 SAVLKS	large surface anti
16	130	93.5	400	1 SAVLVD	large surface anti
17	127	91.4	226	1 JQ1574	major surface anti
18	127	91.4	226	1 SAVLAD	major surface anti
19	127	91.4	226	2 JQ2101	surface antigen -
20	125	89.9	389	2 S20749	surface antigen -
21	124	89.2	226	1 JQ1570	major surface anti
22	124	89.2	226	1 JQ1571	major surface anti
23	124	89.2	226	2 JQ2058	surface antigen -
24	124	89.2	226	2 JQ2075	surface antigen -
25	124	89.2	226	2 JQ2077	surface antigen -
26	124	89.2	226	2 JQ2076	surface antigen -
27	124	89.2	226	2 JQ2061	surface antigen -
28	124	89.2	226	2 JQ2081	surface antigen -
29	124	89.2	226	2 JQ2060	surface antigen -
30	124	89.2	226	2 JQ2079	surface antigen -
31	124	89.2	226	2 JQ2083	surface antigen -
32	124	89.2	226	2 JQ2083	surface antigen -
33	124	89.2	226	2 JQ2062	surface antigen -
34	124	89.2	389	1 SAVLAI	large surface anti
35	124	89.2	389	1 SAVLBH	large surface anti
36	124	89.2	389	1 SAVLJ2	large surface anti
37	124	89.2	389	1 SAVLJ3	large surface anti
38	124	89.2	389	2 S20745	surface antigen -
39	124	89.2	389	2 S20753	surface antigen -
40	124	89.2	389	2 S41871	surface antigen -
41	123	88.5	226	2 JQ2116	surface antigen -
42	122	87.8	226	2 JQ2047	surface antigen -
43	122	87.8	226	2 JQ2051	surface antigen -
44	122	87.8	384	2 T13469	large surface anti
45	122	87.8	389	1 SAVLJ1	large surface anti

ALIGNMENTS

RESULT 1

JQ2054
surface antigen - hepatitis B virus (subtype indet, strain Tar)

C/Species: hepatitis B virus, HBV

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004

C/Accession: JQ2054; PQ0571

R/Norder, H.; Hammes, B. (Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus

J. Gen. Virol. 74, 1341-1348, 1993

A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origi

A/Reference number: JQ2044; MUID:93329382; PMID:8336122

A/Contents: genogroup A

A/Accession: JQ2054

A/Molecule type: DNA

A/Residues: 1-226 <NOR>

A/Cross-references: UNIPROT:Q8JWP7; UNIPROT:Q9Q3D4; UNIPROT:Q8JXH6; UNIPROT:O41759; UNI

PROT:O41755; UNIPROT:O41746; UNIPROT:O90663; UNIPROT:O41744; UNIPROT:Q9JEX46; UNIPROT:O4

JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UNIPROT:Q9W7M6; U

R/Norder, H.; Courouce, A.M.; Magnus, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A/Title: Molecular basis of hepatitis B virus serotype variations within the four major

A/Reference number: PQ0453; MUID:93107848; PMID:1469353

A/Accession: PQ0571

A/Molecule type: DNA

A/Residues: 101-180 <NO2>

A/Cross-references: UNIPARC:UPI0000178434

C/Genetics:

A/Gene: S

C/Keywords: surface antigen

Alignment Scores:	7.82e-12	Length:	226
Pred. No.:	133.00	Matches:	22
Score:	95.65%	Conservative:	0
Percent Similarity:	95.65%	Mismatches:	1
Best Local Similarity:	95.65%	Indels:	0
Query Match:	95.68%	Gaps:	0
DB:	2		

US-10-761-006A-1_COPY_527_595 (1-69) x JQ2054 (1-226)

```
QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProSerAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 2
JQ1577
major surface antigen - hepatitis B virus (subtype ayw1)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JQ1577; JQ2049; PQ0573
R:Norder: H.; Hammas, B.; Loefdahl, S.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 1201-1208, 1992
A:Title: Comparison of the amino acid sequences of nine different serotypes of hepatitis
A:Reference number: JQ1570; MUID:92268879; PMID:1588323
A:Accession: JQ1577
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q69589; UNIPARC:UPI00000F4DF5; GB:X75669; NID:9416076; PIDN:
A:Experimental source: subtype ayw1, strain CNTS-5
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2049
A:Molecule type: DNA
A:Residues: 1-226 <NO2>
A:Cross-references: UNIPARC:UPI00000F4DF5
A:Experimental source: subtype ayw1, strain Mam
R:Norder, H.; Courouce, A.M.; Magnius, L.O.
J. Gen. Virol. 73, 3141-3145, 1992
A:Title: Molecular basis of hepatitis B virus serotype variations within the four major
A:Reference number: PQ0453; MUID:93107848; PMID:1469353
A:Accession: PQ0573
A:Molecule type: DNA
A:Residues: 101-180 <NO3>
A:Cross-references: UNIPARC:UPI000017496F
A:Experimental source: subtype ayw1, strain Mam
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein; surface antigen
F:3/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x JQ1577 (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProSerAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 3
SAVLHV
major surface antigen - hepatitis B virus
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
```

```
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: JQ2093
R:Rivkina, M.B.; Lunin, V.G.; Mahov, A.M.; Tikchonenko, T.I.; Kukain, R.A.
Gene 64, 285-296, 1988
A:Title: Nucleotide sequence of integrated hepatitis B virus DNA and human flanking regi
A:Reference number: JQ2093; MUID:86297159; PMID:2841200
A:Accession: JQ2093
A:Molecule type: DNA
A:Residues: 1-226 <RIV>
A:Cross-references: UNIPROT:P31873; UNIPARC:UPI00001389B6; GB:M21030; NID:9329702; PIDN:
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein; surface antigen
F:3/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x SAVLHV (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProSerAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 4
JQ2052
surface antigen - hepatitis B virus (subtype adw2, strain 8903/86)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2052
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2052
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9W7W6; UNIPARC:UPI00000FA191
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x JQ2052 (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProSerAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 5
```

```
JQ2053
surface antigen - hepatitis B virus (subtype adw2, strain 3138/87)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2053
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2053
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9DLM8; UNIPROT:Q9QAW3; UNIPROT:Q8JWP7; UNIPROT:Q9Q3D4; UNIP
PROT:Q9YVD5; UNIPROT:Q41755; UNIPROT:Q41746; UNIPROT:Q90663; UNIPROT:Q41744; UNIPROT:Q9
WRK7; UNIPROT:Q8JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2053 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAAGCTGC 69
Db 145 GlyAsnCys 147

RESULT 6
JQ2050
surface antigen - hepatitis B virus (subtype adw2, strain 1046/85)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2050
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2050
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q91938; UNIPROT:Q9B6U0; UNIPROT:Q9YVD6; UNIPROT:Q8JWP7; UNIP
PROT:Q9YVD5; UNIPROT:Q41755; UNIPROT:Q41746; UNIPROT:Q90663; UNIPROT:Q41744; UNIPROT:Q9
WRK7; UNIPROT:Q8JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2050 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAAGCTGC 69
Db 145 GlyAsnCys 147

RESULT 6
JQ2050
surface antigen - hepatitis B virus (subtype adw2, strain 1046/85)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2050
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2050
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q91938; UNIPROT:Q9B6U0; UNIPROT:Q9YVD6; UNIPROT:Q8JWP7; UNIP
PROT:Q9YVD5; UNIPROT:Q41755; UNIPROT:Q41746; UNIPROT:Q90663; UNIPROT:Q41744; UNIPROT:Q9
WRK7; UNIPROT:Q8JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2050 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAAGCTGC 69
Db 145 GlyAsnCys 147
```

```
QY 61 GGAAGCTGC 69
Db 145 GlyAsnCys 147

RESULT 7
JQ2048
surface antigen - hepatitis B virus (subtype ayw1, strain Bat)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2048
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2048
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9DLM8; UNIPROT:Q9QAW3; UNIPROT:Q8JWP7; UNIPROT:Q9Q3D4; UNIP
PROT:Q9YVD5; UNIPROT:Q41755; UNIPROT:Q41746; UNIPROT:Q90663; UNIPROT:Q41744; UNIPROT:Q9
WRK7; UNIPROT:Q8JXH0; UNIPROT:Q8JXH3; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UNIPROT:Q91L67; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2048 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAAGCTGC 69
Db 145 GlyAsnCys 147

RESULT 8
JQ2055
surface antigen - hepatitis B virus (subtype adw2, strain B1-85)
C:Species: hepatitis B virus, HBV
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2055
R:Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2055
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81186; UNIPARC:UPI000017842E
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 2,23e-11 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2055 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
```

```
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 9
JQ2078
surface antigen - hepatitis B virus (subtype ayw, strain aywmnt)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2078
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup D
A:Accession: JQ2078
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q67886; UNIPARC:UPI0000178428
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2078 (1-226)

QY      1  ACGACTCCTGCTCAAGGAAACTCTACGTTTCCTCTTGTGCTGACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 10
JQ2046
surface antigen - hepatitis B virus (subtype adw2, strain W1)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2046
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2046
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q91HD6; UNIPROT:Q9Q3C9; UNIPROT:Q72529; UNIPROT:Q91EH7; UNIP
PROT:Q9Q3D4; UNIPROT:Q39886; UNIPROT:Q41754; UNIPROT:Q41758; UNIPROT:Q42029; UNIPROT:Q9Y
1EH8; UNIPROT:Q8JWP5; UNIPROT:Q9WRK7; UNIPROT:Q8JXH0; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2078 (1-226)

QY      1  ACGACTCCTGCTCAAGGAAACTCTACGTTTCCTCTTGTGCTGACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 11
JQ2045
surface antigen - hepatitis B virus (subtype adw2, strain Len)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2045
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2045
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9EGU1; UNIPROT:Q91EH7; UNIPROT:Q8V1B9; UNIPROT:Q9JE50; UNIP
PROT:Q39886; UNIPROT:Q41754; UNIPROT:Q41758; UNIPROT:Q42029; UNIPROT:Q9YVD5; UNIPROT:Q41
1EH8; UNIPROT:Q8JWP5; UNIPROT:Q9WRK7; UNIPROT:Q8JXH0; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2045 (1-226)

QY      1  ACGACTCCTGCTCAAGGAAACTCTACGTTTCCTCTTGTGCTGACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 12
JQ2056
surface antigen - hepatitis B virus (subtype adw2, strain B1-89)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2056
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2056
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81185; UNIPROT:Q81184; UNIPARC:UPI0000178449
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0
```

```
US-10-761-006a-1_COPY_527_595 (1-69) x JQ2046 (1-226)

QY      1  ACGACTCCTGCTCAAGGAAACTCTACGTTTCCTCTTGTGCTGACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 11
JQ2045
surface antigen - hepatitis B virus (subtype adw2, strain Len)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2045
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2045
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9EGU1; UNIPROT:Q91EH7; UNIPROT:Q8V1B9; UNIPROT:Q9JE50; UNIP
PROT:Q39886; UNIPROT:Q41754; UNIPROT:Q41758; UNIPROT:Q42029; UNIPROT:Q9YVD5; UNIPROT:Q41
1EH8; UNIPROT:Q8JWP5; UNIPROT:Q9WRK7; UNIPROT:Q8JXH0; UNIPROT:Q91C42; UNIPROT:Q8JLX2; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2045 (1-226)

QY      1  ACGACTCCTGCTCAAGGAAACTCTACGTTTCCTCTTGTGCTGACAAAACCTTCGGAC 60
      |||
Db      125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 144
      |||
QY      61 GGAAACTGC 69
      |||
Db      145 GlyAsnCys 147

RESULT 12
JQ2056
surface antigen - hepatitis B virus (subtype adw2, strain B1-89)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2056
R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup A
A:Accession: JQ2056
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81185; UNIPROT:Q81184; UNIPARC:UPI0000178449
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      2,23e-11      Length:      226
Score:          130.00      Matches:      21
Percent Similarity: 95.65%      Conservative: 1
Best Local Similarity: 91.30%      Mismatches: 1
Query Match:    93.53%      Indels:      0
DB:             2          Gaps:      0
```

Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ2056 (1-226)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
125 ThrThrProAlaGlnGlyAsnSerMetTyrProSerCysCysThrLysProSerAsp 144
|||||

QY 61 GGAAACTGCG 69
|||||
DB 145 GlyAsnCys 147

RESULT 13
SAVLVE
large surface antigen - hepatitis B virus (subtype adw)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A93460; JQ2044; A03706
R;Ono, Y.; Onda, H.; Sasaki, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.
Nucleic Acids Res. 11, 1747-1757, 1983
A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype
A;Reference number: A93460; MUID:83168919; PMID:6300776
A;Accession: A93460
A;Molecule type: DNA
A;Residues: 1-389 <ON>
A;Cross-references: UNIPROT:P03142; UNIPARC:UPI00001389C4; GB:V00866; GB:J02201
R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A;Reference number: JQ2044; MUID:93329382; PMID:8336122
A;Accession: JQ2044
A;Molecule type: DNA
A;Residues: 164-389 <NOR>
A;Cross-references: UNIPARC:UPI0000156BEO
A;Experimental source: genogroup A, subtype adw2, strain PHBV933
C;Genetics:

A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
P;109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;164-389/Product: major surface antigen (gene S) #status predicted <MSA>
F;4,26,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,01e-11 Length: 389
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x SAVLVE (1-389)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
288 ThrThrProAlaGlnGlyAsnSerLysPheProSerCysCysThrLysProThrAsp 307
|||||

QY 61 GGAAACTGCG 69
|||||
DB 308 GlyAsnCys 310

RESULT 14
JQ1575
major surface antigen - hepatitis B virus
N;Alternate names: envelope protein; HBs antigen
N;Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C;Species: hepatitis B virus, HBV
A;Variety: subtype adw2
C;Date: 31-Dec-1993 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: S47411; JQ1575
R;Plucienniczak, A.

submitted to the EMBL Data Library, August 1994
A;Description: Molecular cloning and sequencing of two complete genomes of polish isolat
A;Reference number: S47404
A;Accession: S47411
A;Molecule type: DNA
A;Residues: 1-400 <PLU>
A;Cross-references: UNIPROT:Q67896; UNIPARC:UPI00000F5BAF; EMBL:Z35717; NID:9527440; PI
R;Norder, H.; Hammas, B.; Loeffel, S.; Courouce, A.M.; Magnus, L.O.
J. Gen. Virol. 73, 1201-1208, 1992
A;Title: Comparison of the amino acid sequences of nine different serotypes of hepatiti
A;Reference number: JQ1570; MUID:92268879; PMID:1588323
A;Accession: JQ1575
A;Molecule type: DNA
A;Residues: 175-400 <NOR>
A;Cross-references: UNIPARC:UPI00000002E4; GB:X75666; NID:9416074; PIDN:CAAS3362.1; PID
A;Experimental source: subtype adw2, strain P6
C;Genetics:

A;Gene: S
A;Introns: 122/3
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
P;1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PSI
F;1-119/Domain: pre-S1 domain #status predicted <PRE1>
F;120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <
F;120-174/Domain: pre-S2 domain #status predicted <PRE2>
F;175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2e-11 Length: 400
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x JQ1575 (1-400)

QY 1 ACGACTCCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
|||||
299 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 318
|||||

Db 61 GGAAACTGCG 69
|||||
Db 319 GlyAsnCys 321

RESULT 15
SAVLVS
large surface antigen - hepatitis B virus (subtype adw, strain 991)
N;Contains: major surface antigen; middle surface antigen
C;Species: hepatitis B virus, HBV
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S10383
R;Koechel, H.G.; Schueler, A.; Lottmann, S.; Thomssen, R.
submitted to the EMBL Data Library, February 1990
A;Reference number: S10380
A;Accession: S10383
A;Molecule type: DNA
A;Residues: 1-400 <KOE>
A;Cross-references: UNIPROT:PI7101; UNIPARC:UPI00001389B8; EMBL:X51970; NID:gil155012; P
C;Genetics:

A;Gene: pre-S1/pre-S2/S
A;Introns: 123/2
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;120-400/Product: middle surface antigen (gene pre-2/S) #status predicted <DSA>
F;175-400/Product: major surface antigen (gene S) #status predicted <MSA>
F;15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2e-11 Length: 400
Score: 130.00 Matches: 21

Percent Similarity: 95.65%
 Best Local Similarity: 91.30%
 Query Match: 93.53%
 DB: 1
 Indels: 0
 Gaps: 0
 Conservative: 1
 Mismatches: 1

US-10-761-006A-1_COPY_527_595 (1-69) x SAVLKS (1-400)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGTGCTGTACAAAACCTTCGGAC 60
 |||||
 Db 299 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 318
 |||||
 QY 61 GGAAACTGC 69
 |||||
 Db 319 GlyAsnCys 321
 |||||

Search completed: December 27, 2005, 21:01:54
 Job time : 4.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:41:12 ; Search time 23.046 Seconds
(without alignments)
4224.720 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595

Perfect score: 139

Sequence: 1 ACAGCTCCCTGCTCAGGAAA.....AACCTTCGACGGAACCTGC 69

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 70528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-Q=/cgn2.1/USPTO.spool_p/US10761006/runat_27122005_192900_15751/app_query.fasta_1.1102
-DB=Uniprot -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPT=0 -LOOPTXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @CGN_1.1.580 @runat_27122005_192900_15751 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	226	Q80H30_HPBVO	Q80h30 hepatitis b
2	139	100.0	226	Q9E8K9_HPBVO	Q9e8k9 hepatitis b
3	139	100.0	400	Q80S15_HPBVO	Q80s15 hepatitis b
4	136	97.8	71	Q76N5_HPBVO	Q76n5 hepatitis b
5	136	97.8	71	Q9YXC2_HPBVO	Q9yxc2 hepatitis b
6	136	97.8	98	Q9QKT6_HPBVO	Q9qkt6 hepatitis b
7	136	97.8	215	Q9Q3D0_HPBVO	Q9q3d0 hepatitis b
8	136	97.8	226	Q8BCB1_HPBVO	Q8bcb1 hepatitis b
9	136	97.8	226	Q9WHS1_HPBVO	Q9whs1 hepatitis b
10	136	97.8	281	Q8BCB2_HPBVO	Q8bcb2 hepatitis b
11	136	97.8	400	Q8BCB3_HPBVO	Q8bcb3 hepatitis b
12	135	97.1	63	Q9Q517_HPBVO	Q9q517 hepatitis b
13	135	97.1	226	Q80BT7_HPBVO	Q80bt7 hepatitis b
14	135	97.1	226	Q9DUJ4_HPBVO	Q9duj4 hepatitis b
15	135	97.1	226	Q9DUK6_HPBVO	Q9duk6 hepatitis b
16	135	97.1	226	Q9E8L3_HPBVO	Q9e8l3 hepatitis b

17	135	97.1	226	2	Q9E8L7_HPBVO	Q9e8l7 hepatitis b
18	135	97.1	400	2	Q39645_HPBVO	Q39645 hepatitis b
19	135	97.1	400	2	Q91525_HPBVO	Q91525 hepatitis b
20	135	97.1	400	2	Q4FDM1_HPBVO	Q4fdm1 hepatitis b
21	134	96.4	47	2	Q9JFX8_HPBVO	Q9jfx8 hepatitis b
22	133	95.7	47	2	Q9JFX9_HPBVO	Q9jfx9 hepatitis b
23	133	95.7	53	2	Q71983_HPBVO	Q71983 hepatitis b
24	133	95.7	53	2	Q71984_HPBVO	Q71984 hepatitis b
25	133	95.7	53	2	Q71987_HPBVO	Q71987 hepatitis b
26	133	95.7	101	2	Q52XM2_HPBVO	Q52xm2 hepatitis b
27	133	95.7	101	2	Q52YEA_HPBVO	Q52ye4 hepatitis b
28	133	95.7	128	2	Q8QVC3_HPBVO	Q8qvc3 hepatitis b
29	133	95.7	128	2	Q8QVC4_HPBVO	Q8qvc4 hepatitis b
30	133	95.7	226	2	Q04258_HPBVO	Q04258 hepatitis b
31	133	95.7	226	2	Q91938_HPBVO	Q91938 hepatitis b
32	133	95.7	226	2	Q5G8Q0_HPBVO	Q5g8q0 hepatitis b
33	133	95.7	226	2	Q76VU2_HPBVO	Q76vu2 duck hepati
34	133	95.7	226	2	Q9Q3C8_HPBVO	Q9q3c8 hepatitis b
35	133	95.7	226	2	Q9Q3E8_HPBVO	Q9q3e8 hepatitis b
36	133	95.7	226	2	Q9YD6_HPBVO	Q9yde6 hepatitis b
37	133	95.7	384	2	Q86986_HPBVO	Q86986 hepatitis b
38	133	95.7	389	2	Q91C55_HPBVO	Q91c55 hepatitis b
39	133	95.7	389	2	Q9YZT8_HPBVO	Q9yzt8 hepatitis b
40	133	95.7	400	2	Q56U13_HPBVO	Q56u13 hepatitis b
41	133	95.7	400	2	Q5Y2B8_HPBVO	Q5y2b8 hepatitis b
42	130	93.5	42	2	Q8QOM2_HPBVO	Q8qom2 hepatitis b
43	130	93.5	42	2	Q8QOM4_HPBVO	Q8qom4 hepatitis b
44	130	93.5	42	2	Q8QOM8_HPBVO	Q8qom8 hepatitis b
45	130	93.5	42	2	Q8QON0_HPBVO	Q8qon0 hepatitis b

ALIGNMENTS

RESULT 1

Q80H30_HPBVO
ID Q80H30_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q80H30; 24, Created
DT 01-JUN-2003 (TREMREL. 24, Last sequence update)
DT 01-JUN-2003 (TREMREL. 24, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthonepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FMU005;
RA Lin X., Zheng D.L., Xu X.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY206376; AAP0547.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR PANTHER; PTHR10832; Hepvir_surfac; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25405 MW; 186C43D50FB8DAA CRC64;

Alignment Scores:
Pred. No.: 1,29e-13 Length: 226
Score: 139.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q80H30_HPBVO (1-226)

QY 1 ACAGCTCCCTGCTCAGGAAACTCTAGCTTTCCCTCTTTGTCGTCAAAACCTTCGGAC 60
|||||
Db 125 ThrThrProAlaGlnGlyAanSerThrPheProSerCysCysThrLysProSerAsp 144
QY 61 GGAACTGC 69

DR Pfam; PF00695; vmsa; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43741 MW; D1DFC54CDF11FA91 CRC64;

Alignment Scores:
Pred. No.: 1.34e-13 Length: 400
Score: 139.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q80515_HPBVO (1-400)

QY 1 ACGACTCTGCTCAAGGAACTCTAGCTTTCCCTCTTGTGTGTACAAAACCTTCGGAC 60
|||||
DB 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318

QY 61 GGAAGACTGC 69
|||||
DB 319 GlyAsnCys 321

RESULT 4
Q767N5_HPBVO
ID Q767N5_HPBVO PRELIMINARY; PRT; 71 AA.
AC Q767N5;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Surface antigen (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lusida M.I., Surayah, Sakugawa H., Nagano-Fujii M., Soetjito,
RA Mulyanto, Handayani R., Boediwarsono, Setiawan P.B., Nidom C.A.,
RA Ogitmoto S., Hotta H.;
RT "Genotype and Subtype Analyses of Hepatitis B Virus (HBV) and Possible
RT Co-infection of HBV and Hepatitis C Virus (HCV) or Hepatitis D Virus
RT (HDV) in Blood Donors, Patients with Chronic Liver Disease and
RT Patients on Hemodialysis in Surabaya, Indonesia.";
RL Microbiol. Immunol. 47:969-975(2003).
DR EMBL; AB113296; BAC99010.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfag.
DR PANTHER; PTHR10832; Hepvir_surfag; 1.
DR Pfam; PF00695; vmsa; 1.
KW Antigen.
FT NON_TER 1
FT NON_TER 71
FT NON_TER 71
SQ SEQUENCE 71 AA; 7861 MW; A8B0FB021386382C CRC64;

Alignment Scores:
Pred. No.: 3.75e-13 Length: 71
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q767N5_HPBVO (1-71)

QY 1 ACGACTCTGCTCAAGGAACTCTAGCTTTCCCTCTTGTGTGTACAAAACCTTCGGAC 60
|||||
DB 13 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 332

QY 61 GGAAGACTGC 69
|||||
DB 33 GlyAsnCys 35

RESULT 5

```

Q9YXC2_HPBVO
ID Q9YXC2_HPBVO PRELIMINARY; PRT; 71 AA.
AC Q9YXC2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface antigen (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98449868; PubMed=9774595;
RA Mayed V.A., Lopez J.B., Telenta P.F.S., Palacios G., Badia I.,
RA Ferro A., Galoppo C., Campos R.;
RT "Distribution of hepatitis B virus genotypes in two different
RT pediatric populations from Argentina.";
RL J. Clin. Microbiol. 36:3362-3365(1998).
DR EMBL; AF043576; AAC79977.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7847 MW; B94A1002139B353C CRC64;

Alignment Scores:
Pred. No.: 3,75e-13 Length: 71
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q9YXC2_HPBVO (1-71)

QY 1 ACAGCTCTCTCAAGAACTCTAGCTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 13 ThrThrProAlaGlnGlyAenSerThrPheProSerCysCysThrLysProThrAsp 32

QY 61 GGAACCTGC 69
Db 33 GlyAenCys 35

RESULT 6
Q9QKT6_HPBVO
ID Q9QKT6_HPBVO PRELIMINARY; PRT; 98 AA.
AC Q9QKT6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HBsAg (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Swenson P.D., Van Geyt C., Alexander R.E., Hagan H.,
RA Freitag-Koontz J.M., De Gendt S., Van Reybroeck G., Wilson S.,
RA Magnus L., Stuyver L.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107159; AAF04986.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER 1
FT NON_TER 98
SQ SEQUENCE 98 AA; 11040 MW; 09B458DE7C0BB075 CRC64;

Alignment Scores:
Pred. No.: 4,04e-13 Length: 215
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q9QKT6_HPBVO (1-215)

QY 1 ACAGCTCTCTCAAGAACTCTAGCTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAenSerThrPheProSerCysCysThrLysProThrAsp 144

QY 61 GGAACCTGC 69
Db 145 GlyAenCys 147

RESULT 8
Q8BCB1_HPBVO
ID Q8BCB1_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q8BCB1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Abn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
DR EMBL; AF537372; AAN15124.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfa; IEA.
DR PANTHER; PTHR10832; Hepvir_surfa; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25433 MW; CBCFB5CAFAA82107 CRC64;

Alignment Scores:
Pred. No.: 4.05e-13 Length: 226
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q8BCB1_HPBVO (1-226)
QY 1 AGCACTCTGCTCAAGAACTCTAGTTTCCCTCTGTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrllysProThrAsp 144

QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 9
Q8BCB2_HPBVO
ID Q8BCB2_HPBVO PRELIMINARY; PRT; 281 AA.
AC Q8BCB2;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Middle S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Abn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
DR EMBL; AF537372; AAN15128.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfa; 1.
DR PANTHER; PTHR10832; Hepvir_surfa; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 281 AA; 31149 MW; 7CA77FAEB18CDBC9 CRC64;

Alignment Scores:
Pred. No.: 4.11e-13 Length: 281
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q8BCB2_HPBVO (1-281)
QY 1 AGCACTCTGCTCAAGAACTCTAGTTTCCCTCTGTGTACAAAACCTTCGGAC 60
Db 180 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrllysProThrAsp 199

QY 61 GGAAACTGC 69
```

```
Db      200 GlyAsnCys 202
|||||
RESULT 11
Q8BCB3 HPBV0
ID Q8BCB3 HPBV0 PRELIMINARY; PRT; 400 AA.
AC Q8BCB3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Large S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22652772; PubMed=12767980;
RX DOI=10.1128/JVI.77.12.6601-6612.2003;
RA Parekh S., Zoulim F., Ahn S.H., Tsai A., Li J., Kawai S., Khan N.,
RA Trepo C., Wands J., Tong S.;
RT "Genome replication, virion secretion, and e antigen expression of
RT naturally occurring hepatitis B virus core promoter mutants.";
RL J. Virol. 77:6601-6612(2003).
RL EMBL; AF537372; AANL5127.1; -; Genomic_DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR PANTHER; PTHR10832; Hepvir_surfac; 1.
DR Pfam; PF00695; vmsa; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43651 MW; 6E6C6630699CC81E CRC64;

Alignment Scores:
Pred. No.: 4.21e-13 Length: 400
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB3 HPBV0 (1-400)

Qy 1 ACGACTCCTGCTCAAGAACTCTAGCTTCCCTTGTGCTGACAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProThrAsp 318
Qy 61 GGAAACTGC 69
Db 319 GlyAsnCys 321

RESULT 12
Q9Q5L7 HPBV0
ID Q9Q5L7 HPBV0 PRELIMINARY; PRT; 63 AA.
AC Q9Q5L7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface antigen (Fragment).
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE
RA Xia G.-L., Nainan O.V., Jia Z.-Y., Cao H.-L., Liu C.-B.,
RA Margolis H.S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF198960; AAF15879.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR PANTHER; PTHR10832; Hepvir_surfac; 1.
DR Pfam; PF00695; vmsa; 1.
KW Antigen.

Alignment Scores:
Pred. No.: 4.21e-13 Length: 400
Score: 136.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.84% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q8BCB3 HPBV0 (1-400)

Qy 1 ACGACTCCTGCTCAAGAACTCTAGCTTCCCTTGTGCTGACAAACCTTCGGAC 60
Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProThrAsp 318
Qy 61 GGAAACTGC 69
Db 319 GlyAsnCys 321

RESULT 13
Q80BT7 HPBV0
ID Q80BT7 HPBV0 PRELIMINARY; PRT; 226 AA.
AC Q80BT7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95154763; PubMed=7851832;
RX Moriyama K., Takada T., Tsutsui Y., Fukada K., Ishibashi H., Niho Y.,
RX Maeda Y.;
RT "Mutations in the transcriptional regulatory region of the precore and
RT core/pregenome of a hepatitis B virus with defective HBeAg
RT production.";
RL Fukuoka Igaku Zasshi 85:314-322(1994).
DR EMBL; S75184; AAP31568.1; -; Genomic_DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR PANTHER; PTHR10832; Hepvir_surfac; 1.
DR Pfam; PF00695; vmsa; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25365 MW; BC8FE3999F34789F CRC64;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x Q80BT7 HPBV0 (1-226)

Qy 1 ACGACTCCTGCTCAAGAACTCTAGCTTCCCTTGTGCTGACAAACCTTCGGAC 60
Db 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144
Qy 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 14
Q9DUJ4 HPBV0
ID Q9DUJ4 HPBV0 PRELIMINARY; PRT; 226 AA.
AC Q9DUJ4;
```

```
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Small surface polypeptide.
GN Name=OG14;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029963; BAB20346.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; vMSA; 1.
KW Antigen; Polypeptide.
SQ SEQUENCE 226 AA; 25375 MW; 7704565C4092E3BC CRC64;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q9DUJ4_HPBVO (1-226)

QY 1 ACGACTCCTGCTCAGGAACTCTACGTTCCCTCTTGTGCTGACAAAACCTTCGGAC 60
Db 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144

QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

Search completed: December 27, 2005, 21:00:25
Job time : 24.046 secs

DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; vMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25186 MW; E57E7DB1CA5DE870 CRC64;

Alignment Scores:
Pred. No.: 5.93e-13 Length: 226
Score: 135.00 Matches: 22
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 0
Query Match: 97.12% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x Q9DUK6_HPBVO (1-226)

QY 1 ACGACTCCTGCTCAGGAACTCTACGTTCCCTCTTGTGCTGACAAAACCTTCGGAC 60
Db 125 ThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 144

QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 15
Q9DUK6_HPBVO
ID Q9DUK6_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9DUK6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Small surface polypeptide.
GN Name=OG2;
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029951; BAB20334.1; -; Genomic_DNA.
DR PIR; JQ2102; JQ2102.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 27, 2005, 20:44:58 ; Search time 45.4 Seconds
(without alignments)
2480.270 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835
Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GTCCTTGGGTATACATTAA 681

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005.192901.15775/app_query.fasta_1.1102
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CIGN 1 101 @runat_27122005.192901.15775 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6
-PGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
1	1230	96.2	400	2	US-09-719-528A-3
2	1230	96.2	400	2	US-10-209-264-3
3	1206	94.4	226	1	US-08-378-011A-3
4	1206	94.4	236	1	US-08-378-011A-1
5	1199	93.8	277	6	5164485-2
6	1190	93.1	843	2	US-09-719-528A-2
7	1190	93.1	843	2	US-10-209-264-2
8	1189	93.0	226	2	US-09-471-573A-2
9	1189	93.0	226	2	US-09-471-573A-40
10	1163	91.0	226	6	5196194-21
11	1163	91.0	226	6	5436139-5
12	1163	91.0	390	2	US-09-721-480-5

13	1163	91.0	531	2	US-09-721-480-7	Sequence 7, Appli
14	1160	90.8	423	1	US-08-760-797A-1	Sequence 1, Appli
15	1160	90.8	424	1	US-08-760-797A-3	Sequence 3, Appli
16	1160	90.8	424	2	US-08-932-929B-1	Sequence 1, Appli
17	1160	90.8	424	2	US-08-932-929B-3	Sequence 3, Appli
18	1157	90.5	281	2	US-09-247-890-10	Sequence 10, Appli
19	1157	90.5	281	2	US-09-724-969-10	Sequence 10, Appli
20	1157	90.5	281	2	US-09-724-852-10	Sequence 10, Appli
21	1156	90.5	281	2	US-09-721-480-3	Sequence 3, Appli
22	1152	90.1	226	6	5198348-1	Patent No. 5198348
23	1151	90.1	225	6	5436139-4	Patent No. 5436139
24	1143	89.5	395	6	5196194-18	Patent No. 5196194
25	1133	88.7	226	2	US-10-104-966-10	Sequence 10, Appli
26	1133	88.7	226	2	US-09-929-955-10	Sequence 10, Appli
27	1126	88.1	226	4	PCT-US96-10602-14	Sequence 14, Appli
28	1126	88.1	281	1	US-08-105-483-214	Sequence 214, App
29	1126	88.1	281	1	US-08-709-209-214	Sequence 214, App
30	1126	88.1	281	1	US-08-458-101-214	Sequence 214, App
31	1126	88.1	281	2	US-09-247-890-12	Sequence 12, Appli
32	1126	88.1	281	2	US-09-724-969-12	Sequence 12, Appli
33	1126	88.1	281	2	US-09-724-852-12	Sequence 12, Appli
34	1126	88.1	389	1	US-08-105-483-216	Sequence 216, App
35	1126	88.1	389	1	US-08-105-483-219	Sequence 219, App
36	1126	88.1	389	1	US-08-709-209-216	Sequence 216, App
37	1126	88.1	389	1	US-08-709-209-219	Sequence 219, App
38	1126	88.1	389	1	US-08-458-101-216	Sequence 216, App
39	1126	88.1	389	1	US-08-458-101-219	Sequence 219, App
40	1126	88.1	389	2	US-08-486-099-106	Sequence 106, App
41	1126	88.1	389	2	US-08-360-107A-116	Sequence 116, App
42	1126	88.1	389	2	US-08-484-223B-106	Sequence 106, App
43	1126	88.1	389	2	US-08-919-597-106	Sequence 106, App
44	1126	88.1	389	2	US-08-475-668A-106	Sequence 106, App
45	1126	88.1	389	2	US-08-485-551A-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-09-719-528A-3
; Sequence 3, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ladas & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/719,528A

FILING DATE: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Mass, Clifford J.

REGISTRATION NUMBER: 30, 086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-719-528A-3

Alignment Scores:
Pred. No.: 3,74e-124 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservativeness: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-719-528A-3 (1-400)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
Db 175 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194

QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGAGTCTCTCTCAAT 119
Db 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214

QY 120 TTTCTAGGGGAGACCCACCGTTCCTCGCCAAAATTCGCGAGTCCCAACCTCCAAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThr-SerAsnHi 234

QY 180 CTCACCAACCTCTGTCCTCCAAATTTGCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db 234 sSerProThrSerCysProProlleCysProGlyTyrArgTrpAsnCysLeuArgGph 254

QY 240 TATCATATTCTCTTCATCTGCTATGCTATGCTCATCTTCTTGTGTCTCTGACTA 299
Db 254 ellellePheLeuPheilleLeuLeuLeuLeuLeuValLeuLeuAspTy 274

QY 300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294

QY 360 ATGCGAGACTCGACACTCTGCTCAAGGAACCTAGCTTTCCCTCTTGTGTCTGTAC 419
Db 294 oCysLysThrCysThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysTh 314

QY 420 AAAACCTTCGGACGGAACTGCACTTGTATTCCCATCCCATCTCTGGGCTTTCGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTATAGGAGTGGCGCTCAGTCCGTTTCTCGGCTCAGTTTACTAGTGCATTCT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 354

QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374

QY 600 TTGGGGGGCAAGTCTGTACAAATCTTGATGCTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 374 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProllePhePh 394

QY 660 TTGCTCTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyIle 400
```

RESULT 2

US-10-209-264-3

; Sequence 3, Application US/10209264

; Patent No. 6787142

; GENERAL INFORMATION:

```
;
; APPLICANT: Oon, Chong Jin
; Lim, Gek Keow
; Zhao, Yi
; Chen, Wei Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
; USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 400 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-209-264-3

Alignment Scores:
Pred. No.: 3,74e-124 Length: 400
Score: 1230.00 Matches: 224
Percent Similarity: 98.68% Conservativeness: 0
Best Local Similarity: 98.68% Mismatches: 2
Query Match: 96.24% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-10-209-264-3 (1-400)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
Db 175 MetGluAsnThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194

QY 61 TCGTTGACAGAATCCTCAATACCGCAGAGTCTAGACTC-TGGTGGAGTCTCTCTCAAT 119
Db 195 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214

QY 120 TTTCTAGGGGAGACCCACCGTTCCTCGCCAAAATTCGCGAGTCCCAACCTCCAAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThr-SerAsnHi 234

QY 180 CTCACCAACCTCTGTCCTCCAAATTTGCTCGCTATCGCTGATGTGTCTGGGGGTTT 239
Db 234 sSerProThrSerCysProProlleCysProGlyTyrArgTrpAsnCysLeuArgGph 254

QY 240 TATCATATTCTCTTCATCTGCTATGCTATGCTCATCTTCTTGTGTCTCTGACTA 299
Db 254 ellellePheLeuPheilleLeuLeuLeuLeuLeuValLeuLeuAspTy 274

QY 300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294

QY 360 ATGCGAGACTCGACACTCTGCTCAAGGAACCTAGCTTTCCCTCTTGTGTCTGTAC 419
Db 294 oCysLysThrCysThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysTh 314

QY 420 AAAACCTTCGGACGGAACTGCACTTGTATTCCCATCCCATCTCTGGGCTTTCGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTATAGGAGTGGCGCTCAGTCCGTTTCTCGGCTCAGTTTACTAGTGCATTCT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 354

QY 540 TCAGTGGTTCGTAGGCTTTCCCACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374

QY 600 TTGGGGGGCAAGTCTGTACAAATCTTGATGCTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 374 rTrpGlyArgSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProllePhePh 394

QY 660 TTGCTCTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyIle 400
```


Db 274 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
Qy 360 ATGCAAGACTGCAGCTCTGCTCAAGAACTCTACGTTCCCTCTCTGTTGCTGTAC 419
Db 294 oCysLeThrCysThrThrProAlaGInGlyAsnSerThrPheProSerCysCysTh 314
Qy 420 AAAACCTTCGCGAGCAACTGCACTTGTATTTCCCATCCCATCATCTCTGGGCTTTGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 334
Qy 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCATTGT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
Qy 540 TCAGTGTTCGTAGGCTTTCCCACTGTTTGGCTTTCAGTTATATGATGATGCTA 599
Db 354 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 374
Qy 600 TTGGGGCGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 374 rTrpGlyArgSerLeuLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
Qy 660 TTGCTTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTrpIle 400

RESULT 3

US-08-378-011A-3
; Sequence 3, Application US/08378011A
; Patent No. 5693497
; GENERAL INFORMATION:
; APPLICANT: TAKAMIZAWA, Akihisa
; APPLICANT: FUJITA, Hiroyuki
; APPLICANT: MANABE, Sadao
; APPLICANT: KATO, Masahiko
; APPLICANT: OSAME, Juichiro
; APPLICANT: YOSHIDA, Iwao
; APPLICANT: KONOBE, Takeo
; APPLICANT: TAKAKU, Keisuke
; TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND A PROCESS FOR
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII, generated using Word Perfect, version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,011A
; FILING DATE: 25-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 61-143412
; FILING DATE: 18-JUN-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/061,518
; FILING DATE: 15-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,494
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 870602B
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-011A-3
Alignment Scores:
Pred. No.: 1,18e-121 Length: 226
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-08-378-011A-3 (1-226)

Qy 1 ATGAGAACACAACTCAGATTCTAGGACCCCTCTCTGTGTACAGGGGGTTTTTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
Qy 61 TCGTTGACAAGATCCTCACAATACCGCAGAGCTAGACTG-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
Qy 120 TTTCTAGGGGAGCAGCCAGCTGTTCTGGCCAAAATTCAGTCCCAACCTCCCAATCA 179
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
Qy 180 CTCACCAACTTGTCTCTCAATTTGTCTGCTATCGCTGGATGCTCGGGCGTTT 239
Db 60 sSerProThrSerCysProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
Qy 240 TATCATATTCCTCTCTCATCTGCTGCTATGCCCTCATCTTCTTGTGGTCTTCTGGACTA 299
Db 80 eIleIlePhePheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
Qy 300 CCAAGGTATGTTGCCGTTTGTCTCTTACTTCCAGGAACATCAACACACAGCAGGGGCC 359
Db 100 rGInGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
Qy 360 ATGCAAGACTGCAGCTCTCTCTCAAGAACTCTAGCTTTCCCTCTTGTGCTGTAC 419
Db 120 oCysLeThrCysThrIleProAlaGInGlyThrSerMetPheProSerCysCysTh 140
Qy 420 AAAACCTTCGCGAGCAACTGCACTTGTATTTCCATCCCATCATCTCTGGGCTTTCGCAAG 479
Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
Qy 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCATTGT 539
Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
Qy 540 TCAGTGTTCGTAGGCTTTCCCACTGTTTGGCTTTCAGTTATATGATGATGCTA 599
Db 180 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 200
Qy 600 TTGGGGCGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTrpAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
Qy 660 TTGCTTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTrpIle 226

RESULT 4

US-08-378-011A-1
; Sequence 1, Application US/08378011A
; Patent No. 5693497

```

; GENERAL INFORMATION:
; APPLICANT: TAKAMIZAWA, Akihisa
; APPLICANT: FUJITA, Hiroyuki
; APPLICANT: MANABE, Sadao
; APPLICANT: KATO, Masahiko
; APPLICANT: OSAME, Juichiro
; APPLICANT: YOSHIDA, Iwao
; APPLICANT: KONOE, Takeo
; APPLICANT: TAKAKU, Keisuke
; TITLE OF INVENTION: HEPATITIS B VIRUS ANTIGEN AND A PROCESS FOR
; TITLE OF INVENTION: PRODUCING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McIeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII, generated using Word Perfect, version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,011A
; FILING DATE: 25-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 61-143412
; FILING DATE: 18-JUN-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/061,518
; FILING DATE: 15-JUN-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/902,494
; FILING DATE: 23-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 870602B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-011A-1

Alignment Scores:
Pred. No.: 1,2e-121 Length: 236
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x US-08-378-011A-1 (1-236)

QY 1 ATGGAGAACACATCATCAGATTCCTAGGACCCCTGCTGTATACAGCGGGGTTTTC 60
Db 11 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 30
QY 61 TCGTTTACAAAGATCCTCAATACCGCAGAGCTAGATCTC-TGGTGGACTTCTCTCAAT 119
Db 31 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 50
QY 120 TTTCTAGGGGGAGCACCCACGCTGTTCTCTGCCAAAATTCGAGTCCCCCAACCTCCAATCA 179

```

```

Db 51 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 70
QY 180 CTCACCAACCTCTGTCTCCTCAATTTGTCCTGGGTATCGCTGGATGCTCTCGCGGCTTT 239
Db 70 sSerProThrSerCysProPheLeuCysProGlyYrArgTrpMetCysLeuArgArph 90
QY 240 TATCATATTCCTCTTCATCTCGCTGCTATGCTCATCTCTTGTGGTCTTCTTGACTA 299
Db 90 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspfy 110
QY 300 CCAAGGTATGTTGCCGTTTGTCTCTACTTCCAGGAACATCAACACCAGCAGCGGGCC 359
Db 110 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 130
QY 360 ATGCAAGACCTGCGACGACTCTCTGCTCAAGGAAACTCTACGTTCCCTCTTGTGTCTGTAC 419
Db 130 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 150
QY 420 AAACCTTTCGGACCGGAAACTGCACCTTGTATTCCCATCCCATCATCTCGGCTTTCGCAAG 479
Db 150 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 170
QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTCCCAATTTGT 539
Db 170 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 190
QY 540 TCAGTGGTTCGTAGGGCTTCCCCACACTGTTGGCTTTCAGTTATATGATGATGTGTA 599
Db 190 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 210
QY 600 TTGGGGCGGAAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTTCT 659
Db 210 rTrpGlyProSerLeuTyraIleLeuSerProPheLeuProLeuLeuProIlePheph 230
QY 660 TTGCTTTGGGTATACATT 678
Db 230 eCysLeuTrpValTyrlle 236

RESULT 5
5164485-2
; Patent No. 5164485
; APPLICANT: FUJISAWA, YUKIO; ITOH, YASUAKI; NISHIMURA, OSAMU
; FUJII, TOMOKO
; TITLE OF INVENTION: MODIFIED HEPATITIS B VIRUS SURFACE
; ANTIGEN P31 AND PRODUCTION THEREOF
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/547,948
; FILING DATE: 03-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 898,425
; FILING DATE: 20-AUG-1986
; SEQ ID NO:2:
; LENGTH: 277
; 5164485-2

Alignment Scores:
Pred. No.: 7,27e-121 Length: 277
Score: 1199.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 93.82% Indels: 2
DB: 6 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x 5164485-2 (1-277)

QY 1 ATGGAGAACACATCATCAGATTCCTAGGACCCCTGCTGTATACAGCGGGGTTTTC 60
Db 52 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 71
QY 61 TCGTTTACAAAGATCCTCAATACCGCAGAGCTAGATCTC-TGGTGGACTTCTCTCAAT 119

```

```

Db 72 LeuLeuThrArglleLeuThrilleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 91
Qy 120 TTTCTAGGGGAGCACCACGCTGCTCTGCGCCAAAATTCGACGTCCCAACCTCCCAATCA 179
Db 92 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 111
Qy 180 CTCACCAACCTCTGCTCTCAATTTGCTCTGCTATCGGTATCGGTATGCTGGGGCTTT 239
Db 111 sSerProThrSerCysProThrCysProGlyGlnAsnSerGlnSerProThrSerAsnHi 131
Qy 240 TATCATATCTCTTCACTGCTGCTGCTATGCTGCTATGCTGCTGCTGCTGCTGCTGCT 299
Db 131 elleillePheLeuPheilleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 151
Qy 300 CCAAGGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
Db 151 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 171
Qy 360 ATGCAAGACCTGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Db 171 oCysLeuThrCysThrilleProAlaGlnGlyThrSerMetPheProSerCysCysTh 191
Qy 420 AAAACCTTCGACGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 191 rLysProSerAspGlyAsnCysThrCysilleProilleProSerSerTrpAlaPheAla 211
Qy 480 ATTCTATGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Db 211 gPheLeuTrpLysAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 231
Qy 540 TCAGTGTGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
Db 231 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValilleTrpMetMetTrp 251
Qy 600 TTGGGGCGAGTCTGACACACTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 659
Db 251 rTrpGlyProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 271
Qy 660 TTGCTCTTTGGGTATACATT 678
Db 271 eCysLeuTrpVallyrille 277

```

RESULT 6

US-09-719-528A-2
 ✓ Sequence 2, Application US/09719528A
 Patent No. 6558626

GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin
 Lim, Gek Keow
 Zhao, Yi
 Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

USSES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ladas & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

APPLICATION DATA:

CURRENT APPLICATION NUMBER: US/09/719,528A

FILING DATE: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:
 NAME: Mass, Clifford J.
 REGISTRATION NUMBER: 30,086
 REFERENCE/DOCKET NUMBER: U-013109-7
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 708-1800
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 843 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-719-528A-2

Alignment Scores:
 Pred. No.: 1,04e-119 Length: 843
 Score: 1190.00 Matches: 226
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 93.11% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-719-528A-2 (1-843)

```

Qy 3 GGAAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 62
Db 356 GlyGluHisAsnIleArgIleProArgThrProAlaArgValThrGlyGlyValPheLeu 375
Qy 63 GTTGACAAGAATCTCACAAATACCGCAGAGTCTAGACTCTGGTGGACTCTCTCAATTTT 122
Db 376 ValAspLysAsnProHisAsnThrAlaGluSerArgLeuTrpTrpThrSerLeuAsnPh 395
Qy 123 CTAGGGGGGACGACCCACGCTGTTCTGCGCCAAAATTCGACGTCCCAACCTCCCAATCAC 182
Db 396 LeuGlyGlyAlaProThrCysSerTrpProLysPheAlaValProAsnLeuGlnSerLeu 415
Qy 183 ACCAACTCTCTGCTCCCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
Db 416 ThrAsnLeuLeuSerSerAsnLeuSerTrpLeuSerLeuAspValSerAlaAlaPheTyr 435
Qy 243 CATATTCTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Db 436 HisIleProLeuHisProAlaAlaMetProHisLeuLeuValGlySerSerGlyLeuPro 455
Qy 303 AGGTATGTTGCGCGTTTGTCTCTACTTCCAGGAACATCAACACCAGCAGCAGCGGGCCAT 362
Db 456 ArgTyrValAlaArgLeuSerSerThrSerArgAsnIleAsnHisGlnHisGlyAlaMet 475
Qy 363 CAAGACTGCGACGACTCTGCTGCTCAAGAAACTCTAGCTTTCCCTCTTGTGCTGCTGCT 422
Db 476 GlnAspLeuHisAspSerCysSerArgLysLeuTyrValSerLeuLeuLeuLeuTyrLys 495
Qy 423 ACCTTCGGACGGAACCTGCACTTGTATTCCCATCCCATCATCTGCTGCTGCTGCTGCTGCT 482
Db 496 ThrPheGlyArgLysLeuHisLeuTyrSerHisProIleIleLeuGlyPheArgLysIle 515
Qy 483 CCTATGGAGTGGGCTCAGTCCGTTTCTCCTGGCTCAGTTTACTAGTGCATTTGTTTCA 542
Db 516 ProMetGlyValGlyLeuSerProPheLeuLeuAlaGlnPheThrSerAlaIleCysSer 535
Qy 543 GTGTTCTGTTAGGGCTTTCCCCCACTGTTTGGCTTTTCCAGTTATATGAGTATGTTGTT 602
Db 536 ValValArgArgAlaPheProHisCysLeuAlaPheSerTyrMetAspValValLeu 555
Qy 603 GGGCGGAGTCTGCTACACATCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTTTTG 662
Db 556 GlyAlaLysSerValGlnHisLeuGluSerLeuPheThrSerIleThrAsnPheLeuLeu 575
Qy 663 TCTTTGGGTATACATTTA 680
Db 576 SerLeuGlyIleHisLeu 581

```

```
RESULT 7
US-10-209-264-2
; Sequence 2, Application US/10209264
; Patent No. 6787142
; GENERAL INFORMATION:
; APPLICANT: Oon, Chong Jin
;           Lim, Gek Keow
;           Zhao, Yi Ning
; TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
;           USES THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/209,264
; FILING DATE: 31-Jul-2002
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SG98/00046
; FILING DATE: 19-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-013109-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-209-264-2
Alignment Scores:
Pred. No.: 1,04e-119 Length: 843
Score: 1190.00 Matches: 226
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.11% Indels: 0
DB: 2 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x US-10-209-264-2 (1-843)
QY 3 GGAGAACACAAATCAGGATTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTCTC 62
Db 356 GlyGluHisAsnIleArgIleProArgThrProAlaArgValThrGlyGlyValPheLeu 375
QY 63 GTTGACAGAGATCCTCACAATACCGCAGAGTCTAGACTCTGGTGGACTTCTCTCAATTTT 122
Db 376 ValAspLysAsnProHisAsnThrAlaGluSerArgLeuTrpTrpThrSerLeuAsnPhe 395
QY 123 CTAGGGGGAGCACCACGCTGTTCTCGCCAAAATTCGCAGTCCCAACCTCCAATCACATC 182
Db 396 LeuGlyGlyAlaProThrCysSerTrpProLysPheAlaValProAsnLeuGlnSerLeu 415
QY 183 ACCAACCTCTGTCTCCAAATTTGTCCTGCTATCGCTGGAGTGTCTGGCGGGTTTTTAT 242
Db 416 ThrAsnLeuLeuSerSerAsnLeuSerTrpLeuSerLeuAspValSerAlaAlaPheTyr 435
QY 243 CATATTCCTTTCATCTCTGCTGCTATGCCTCATCTTCTTGTGGTCTTCTGGACTACCA 302
```

```
Db 436 HisIleProLeuHisProAlaAlaMetProHisLeuLeuValGlySerSerGlyLeuPro 455
QY 303 AGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACCAACAGCAGCGGGCCCATG 362
Db 456 ArgTyrValAlaArgLeuSerSerThrSerArgAsnIleAsnHisGlnHisGlyAlaMet 475
QY 363 CAAGACCTGCAGACTCCTCGTCTCAAGAAACTCTACGTTTCCCTTGTGCTGTGTACAAA 422
Db 476 GlnAspLeuHisAspSerCysSerArgLysLeuTyrValSerLeuLeuLeuTyrLys 495
QY 423 ACCTTCGACGAGAAACTGCACCTTGATTCCCATCCATCATCTCGGCTTTTCGCAAGATT 482
Db 496 ThrPheGlyArgLysLeuHisLeuTyrSerHisProIleLeuGlyPheArgLysIle 515
QY 483 CCTATGGAGTGGGCCTCAGTCCCGTTTCTCCTGGCTCAGTTTACTAGTCCCATTTGTTC 542
Db 516 ProMetGlyValGlyLeuSerProPheLeuLeuAlaGlnPheThrSerAlaIleCysSer 535
QY 543 GTGGTTCGTAGGCTTTCCCCACATGTTTGGCTTTTCACTATATGATGATGATGATGATTG 602
Db 536 ValValArgArgAlaPheProHisCysLeuAlaPheSerTyrMetAspValValLeu 555
QY 603 GGGCGAAGTCTGTACAACATCTTGAGTCCCTTTTACTCTCTATTACCAATTTCTCTTTG 662
Db 556 GlyAlaLysSerValGlnHisLeuGluSerLeuPheThrSerIleThrAsnPheLeuLeu 575
QY 663 TCTTTGGGTATACATTTA 680
Db 576 SerLeuGlyIleHisLeu 581
RESULT 8
US-09-471-573A-2
; Sequence 2, Application US/09471573A
; Patent No. 6551820
; GENERAL INFORMATION:
; APPLICANT: Mason, Hugh
; TITLE OF INVENTION: Expression of Immunogenic Hepatitis B Surface Antigens In Transge
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 3121/1080
; CURRENT APPLICATION NUMBER: US/09/471,573A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,827
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Wild-type HBeAg amino acid sequence
; NAME/KEY: misc_feature
; OTHER INFORMATION: Wild-type HBeAg amino acid sequence
US-09-471-573A-2
Alignment Scores:
Pred. No.: 8,1e-120 Length: 226
Score: 1189.00 Matches: 218
Percent Similarity: 96.04% Conservative: 0
Best Local Similarity: 96.04% Mismatches: 8
Query Match: 93.04% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x US-09-471-573A-2 (1-226)
QY 1 ATGGAGAACACAAATCAGGATTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTCTC 60
Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAGATCCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
```



```
Alignment Scores:
Pred. No.: 5,22e-117 Length: 226
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 5436139-5 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x 5196194-21 (1-226)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db 1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGACACCCAGGTGTTCTGCGCAAAATTCGAGTCCCAACCTCCAATCA 179
Db 41 PheLeuGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTGTCTCCAAATTTGCTGCTGCTATCGCTGGATGCTGCGGGGTTT 239
Db 60 sSerProThrSerCysProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTGCTCATCTTCTGTTGTTCTCTGACTA 299
Db 80 eillelePheLeuPheleleuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGACGGAACCTGCTATTCCTCCATCCATCCTCGGCTTTCGCAAG 479
Db 140 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
QY 480 ATTCTATGGAGTGGCTGTGACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 539
Db 160 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
QY 540 TCAGTGGTTCGTAGGCTTCCCACTGTTGCTTTCAGTTATATGATGATGATGATA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACAACTCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTrpSerIleValSerProPheIleProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226
```

RESULT 11

```
5436139-5
; Patent No. 5436139
; APPLICANT: RUTTER, WILLIAM J.;GOODMAN, HOWARD M.
; TITLE OF INVENTION: NON-PASSAGEABLE VIRUS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/058/89,993
; FILING DATE: 08-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 571,331
; FILING DATE: 22-AUG-1990
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
```

```
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO:5:
; LENGTH: 226
5436139-5

Alignment Scores:
Pred. No.: 5,22e-117 Length: 226
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 5436139-5 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x 5436139-5 (1-226)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTTACAGCGGGGTTTTTC 60
Db 1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCGTTGACAGAATCCTCACAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
QY 120 TTTCTAGGGGAGACACCCAGGTGTTCTGCGCAAAATTCGCAAGTCCCAACCTCCAATCA 179
Db 41 PheLeuGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 60
QY 180 CTCACCAACCTCTGTCTCCAAATTTGCTGCTGCTATCGCTGGATGCTGCGGGGTTT 239
Db 60 sSerProThrSerCysProIleCysProGlyTrpArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTGCTCATCTTCTTGTGTTCTCTGACTA 299
Db 80 eillelePheLeuPheleleuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTCGACGGAACCTGCTATTCCTCCATCCATCCTCGGCTTTCGCAAG 479
Db 100 rGlnGlyMetLeuProValCysProLeuIleProGlySerTrpThrSerThrGlyPr 120
QY 360 ATGCAAGACCTGACGACTCTGCTGCTCAAGGAACTCTAGTTCCCTCTTGTGCTGTAC 419
Db 120 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 140
QY 420 AAAACCTTCGACGGAACCTGACACTGTCACATTCCTCCATCCATCCTCGGCTTTCGCAAG 479
Db 140 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
QY 480 ATTCTATGGAGTGGCTGTGACAACTCTGAGTCCCTTTTACCTCTATTACCAATTTCTT 539
Db 160 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
QY 540 TCAGTGGTTCGTAGGCTTCCCACTGTTGCTTTCAGTTATATGATGATGATGATA 599
Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGGAGTCTGTACAACTCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTrpSerIleValSerProPheIleProLeuLeuProIlePhePh 220
QY 660 TTGCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226

RESULT 12
US-09-721-480-5
; Sequence 5, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
```

```
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PPO1635.002
; CURRENT APPLICATION NUMBER: US/09/721.480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCMVII opti
; OTHER INFORMATION: 330 Ei/SAG
US-09-721-480-5

Alignment Scores:
Pred. No.: 6.44e-117 Length: 390
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-721-480-5 (1-390)

Qy 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTGTGTACAGGCGGGTTTTTC 60
Db 165 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 184
Qy 61 TCGTTGACAGAAATCCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 185 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 204
Qy 120 TTTCTAGGGGAGACACCCAGCTGTCTCGGCCAAATTCGAGTCCCAACCTCCAATCA 179
Db 205 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnH1 224
Qy 180 CTCACCAACCTCTTGTCTCCAAATTTGTCTGGCTATCGCTGGAGTGTCTGCGCGCTTT 239
Db 224 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgGph 244
Qy 240 TATCATATCTCTTCATCCTGCTGCTATGCTCATCTTCTTGTGTGTCTTCTGGACTA 299
Db 244 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 264
Qy 300 CCAAGGTATGTGCGCGTTTGTCTCTACTTCCAGGAACATCAACCCAGCAGCGGGCC 359
Db 264 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrThrSerThrGlyPr 284
Qy 360 ATGCAAGACCTGCGACTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db 284 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 304
Qy 420 AAAACCTTCGAGCGAAACTGCACTGTATTCCTCCATCCCATCATCCTGGGCTTCGCAAG 479
Db 304 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 324
Qy 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTGCTCAGTTTACTAGTGCATTTGT 539
Db 324 sTyrLeuTrpGlnTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 344
Qy 540 TCAGTGGTTCGTAGGGCTTTCCCCACCTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db 344 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 364
Qy 600 TTGGGGCGAAGTCTGTACAACTCTGAGTCCCTTTTACTCTATTACCAATTTCTT 659
Db 364 rTrpGlyProSerLeuTyrSerIleValSerProPheIleProLeuLeuProIlePhePhe 384
Qy 660 TTGCTTTTGGGTATACATT 678
Db 384 eCysLeuTrpValTyrIle 390
```

```
RESULT 13
US-09-721-480-7
; Sequence 7, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PPO1635.002
; CURRENT APPLICATION NUMBER: US/09/721.480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pCMV-II-E2661-sag
US-09-721-480-7

Alignment Scores:
Pred. No.: 7.24e-117 Length: 531
Score: 1163.00 Matches: 208
Percent Similarity: 96.04% Conservative: 10
Best Local Similarity: 91.63% Mismatches: 8
Query Match: 91.00% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-09-721-480-7 (1-531)

Qy 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTGTGTACAGGCGGGTTTTTC 60
Db 306 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 325
Qy 61 TCGTTGACAGAAATCCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 326 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 345
Qy 120 TTTCTAGGGGAGACACCCAGCTGTCTCGGCCAAATTCGAGTCCCAACCTCCAATCA 179
Db 346 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnH1 365
Qy 180 CTCACCAACCTCTTGTCTCCAAATTTGTCTGGCTATCGCTGGAGTGTCTGCGCGCTTT 239
Db 365 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgGph 385
Qy 240 TATCATATCTCTTCATCCTGCTGCTATGCTCATCTTCTTGTGTGTGTCTTCTGGACTA 299
Db 385 eIleIlePheLeuPheIleLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 405
Qy 300 CCAAGGTATGTGCGCGTTTGTCTCTACTTCCAGGAACATCAACCCAGCAGCGGGCC 359
Db 405 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrThrSerThrGlyPr 425
Qy 360 ATGCAAGACCTGCGACTCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db 425 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 445
Qy 420 AAAACCTTCGAGCGAAACTGCACTGTATTCCTCCATCCCATCATCCTGGGCTTCGCAAG 479
Db 445 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 465
Qy 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTGCTCAGTTTACTAGTGCATTTGT 539
Db 465 sTyrLeuTrpGlnTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 485
Qy 540 TCAGTGGTTCGTAGGGCTTTCCCCACCTGTTTGGCTTTTCAGTTATATGATGATGTGTA 599
Db 485 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 505
```

```

QY 600 TTGGGGCGAGTCTGTACACACTTGTAGTCCCTTTTACCTCTATTACCAATTTCCTT 659
Db 505 rTrpGlyProSerLeuTy-SerIleValSerProPheIleProLeuLeuProIlePhePh 525
QY 660 TTGTCTTTGGGTATACATT 678
Db 525 eCysLeuTrpValTyrIle 531

RESULT 14
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-760-797A-1

Alignment Scores:
Pred. No.: 1.4e-116 Length: 423
Score: 1160.00 Matches: 207
Percent Similarity: 96.04% Conservative: 11
Best Local Similarity: 91.19% Mismatches: 8
Query Match: 90.77% Indels: 2
DB: Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x US-08-760-797A-1 (1-423)

QY 1 ATGGAGAACACATCAGATTCTAGGACCCCTGCTGTGTACAGGGGGTTC 60
Db 198 MetGluAsnIleThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePh 217
QY 61 TCGTTGACAGAAATCTCAATACCCGACAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 218 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 237
QY 120 TTTCTAGGGGAGACCCAGCTGTTCTGGCGAAATTCGAGTCCCGAACCTCCCAATCA 179

```

```

Db 238 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 257
QY 180 CTCACCAACCTCTTGTCTCCTCCAAATTTGTCCTGGGTATCGCTGGATGTCTCTCGCGCGTTT 239
Db 257 sSerProThrSerCysProPheIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 277
QY 240 TATCATATTCTCTTCATCCTGCTGCTGCTATGCTCATCTTCTTGTGGTGTCTTCGACTA 299
Db 277 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 297
QY 300 CCAAGGTATGTTGCCGTTTGTCTCTACTTCCAGAAACATCAACACACAGACGCGGCC 359
Db 297 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrAsnThrGlyPr 317
QY 360 ATGCAAGACCTGCGACGACTCTCTCAAGGAAACTCTACGTTTCCCTCTTGTGTCTGTAC 419
Db 317 oCysIleThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysTh 337
QY 420 AAAACCTTCGGACGGAAACTGCTGATTTGATTCCTCCATCCCATCCTGGGCTTCGCAAG 479
Db 337 rIysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaTy 357
QY 480 ATTCCTATGGGAGTGGGCTCAGTCCGTTTCTCCTGGCTCAGTTTACTAGTCCCAATTTGT 539
Db 357 sTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 377
QY 540 TCAGTGGTTCGTAGGGCTTCCCCCAGTGTGTTGGCTTTTTCAGTTATATGATGATGTGTA 599
Db 377 GlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 397
QY 600 TTGGGGCGAGTCTGTACACACTTGTGACATCTTGTAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db 397 rTrpGlyProSerLeuTy-SerIleValSerProPheIleProLeuLeuProIlePhePh 417
QY 660 TTGTCTTTGGGTATACATT 678
Db 417 eCysLeuTrpValTyrIle 423

RESULT 15
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096

```


; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-3

Alignment Scores:
Pred. No.: 1.4e-116 Length: 424
Score: 1160.00 Matches: 207
Percent Similarity: 96.04% Conservative: 11
Best Local Similarity: 91.19% Mismatches: 8
Query Match: 90.77% Indels: 2
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x US-08-760-797A-3 (1-424)

```
QY 1 ATGGAGAACAAACATCAGAGATTCCTAGGACCCCTGCTCGTGTTCACAGCGGGGTTTTTC 60
DB 199 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 218
QY 61 TCGTTGACAGAACTCTCACAATACCGCAGAGCTAGACTC-TGGTCGAGCTTCTCTCAAT 119
DB 219 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 238
QY 120 TTTCTAGGGGAGCACCACCGTGTCTCTGCGCCAAATTCGCAGTCCCAACCTCCAATCA 179
DB 239 PheLeuGlyGlySerProValCysLeu-GlyGlnAsnSerGlnSerProThrSerAsnHi 258
QY 180 CTCACCAACCTTTGTCTCAATTTGCTCGGCTATCGTGGATGTCGTGGCGGCTTT 239
DB 258 sSerProThrSerCysProPheCysProGlyTyArgTrpMetCysLeuArgArgPh 278
QY 240 TATCATATTCTCTTCATCTCTGCTGCTATGCTCATCTTCTTGTGGTCTTCTGGACTA 299
DB 278 eIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 298
QY 300 CCAAGGTATGTTGGCGGTTTGTCTCTCTACTTCCAGGAACATCAACACCGCGGGGCC 359
DB 298 rGlnGlyMetLeuProValCysProLeuIleProGlySerThrThrAsnThrGlyPr 318
QY 360 ATGCAAGACCTGCACGACTCTGCTCAAGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
DB 318 oCysLysThrCysThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysTh 338
QY 420 AAAACCTTCGGACGGAACCTGCACTTGTATTCCTCCATCCCATCCTGGGCTTTCGCAAG 479
DB 338 rLysProThrAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 358
QY 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGCTCAGTTTACTAGTGCCTATTTGT 539
DB 358 sTyLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 378
QY 540 TCAGTGTTCCTAGGGCTTTCCTCCACTGTTTGGCTTTCAGTTATATGATGATGTGCTA 599
DB 378 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerAlaIleTrpMetMetTrpTy 398
QY 600 TTGGGGGGGAAGTCTGTACAACATCTTGAGTCCCTTTTACCTCTATATACCAATTTCTT 659
DB 398 rTrpGlyProSerLeuTySerIleValSerProPheIleProLeuLeuProIlePhePh 418
QY 660 TTGTCTTTGGGTATACATT 678
DB 418 eCysLeuTrpValIle 424
```

Search completed: December 27, 2005, 21:03:41
Job time : 52.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

QM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:41:12 ; Search time 227.454 Seconds
(without alignments)
4224.720 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835

Perfect score: 1278

Sequence: 1 ATCGAGACACACATCAGG.....GTCTTGGGTATACATTAA 681

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192900_15751/app_query.fasta_1.1102
-DB-Uniprot -OPMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USERS=US10761006 @CNG 1 1 580 @runat_27122005_192900_15751 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1220	95.5	226	2 Q9DUJ4_HPBVO	Q9duj4 hepatitis b
2	1213	94.9	400	2 Q7TDS6_HPBVO	Q7tds6 hepatitis b
3	1212	94.8	226	2 Q9DHA2_HPBVO	Q9dha2 hepatitis b
4	1210	94.7	226	2 Q80GU3_HPBVO	Q80gu3 hepatitis b
5	1210	94.7	400	2 Q80GU5_HPBVO	Q80gu5 hepatitis b
6	1209	94.6	400	2 Q4FD52_HPBVO	Q4fd52 hepatitis b
7	1208	94.5	400	2 Q5SDK8_HPBVO	Q5sdk8 hepatitis b
8	1208	94.5	400	2 Q91GW8_HPBVO	Q91gw8 hepatitis b
9	1207	94.4	226	2 Q9PX13_HPBVO	Q9px13 hepatitis b
10	1207	94.4	392	2 Q9QC26_HPBVR	Q9qc26 hepatitis b
11	1206	94.4	226	2 Q9W966_HPBVO	Q9w966 hepatitis b
12	1206	94.4	226	2 Q9YK19_HPBVO	Q9yk19 hepatitis b
13	1206	94.4	399	1 VMSA_HPBVR	P03140 hepatitis b
14	1206	94.4	400	2 Q68RP7_HPBVO	Q68rp7 hepatitis b
15	1206	94.4	400	2 Q7TDR4_HPBVO	Q7tdr4 hepatitis b
16	1206	94.4	400	2 Q80MQ7_HPBVO	Q80mq7 hepatitis b

17	1206	94.4	400	2 Q8VIM3_HPBVO	Q8vim3 hepatitis b
18	1206	94.4	400	2 Q91GW7_HPBVO	Q91gw7 hepatitis b
19	1206	94.4	400	2 Q9YZT1_HPBVO	Q9yzt1 hepatitis b
20	1205	94.3	226	2 Q9DUJ3_HPBVO	Q9duj3 hepatitis b
21	1204	94.2	226	2 Q9DUJ6_HPBVO	Q9duj6 hepatitis b
22	1204	94.2	226	2 Q9Q3D1_HPBVO	Q9q3d1 hepatitis b
23	1204	94.2	226	2 Q9QT41_HPBVO	Q9qt41 hepatitis b
24	1204	94.2	400	2 Q91590_HPBVO	Q91590 hepatitis b
25	1204	94.2	400	2 Q7TDR2_HPBVO	Q7tdr2 hepatitis b
26	1203	94.1	226	2 Q5KR14_HPBVO	Q5kr14 hepatitis b
27	1203	94.1	226	2 Q9DHB0_HPBVO	Q9dho hepatitis b
28	1203	94.1	226	2 Q9DHB0_HPBVO	Q9dho hepatitis b
29	1203	94.1	226	2 Q9PX64_HPBVO	Q9px64 hepatitis b
30	1203	94.1	281	2 Q80GU4_HPBVO	Q80gu4 hepatitis b
31	1203	94.1	400	2 Q913B0_HPBVO	Q913b0 hepatitis b
32	1203	94.1	400	2 Q9DTC8_HPBVO	Q9dte8 hepatitis b
33	1202	94.1	226	2 Q80GX9_HPBVO	Q80gx9 hepatitis b
34	1202	94.1	226	2 Q9DH82_HPBVO	Q9dh82 hepatitis b
35	1202	94.1	226	2 Q9DUJ1_HPBVO	Q9duj1 hepatitis b
36	1202	94.1	226	2 Q9QT46_HPBVO	Q9qt46 hepatitis b
37	1202	94.1	226	2 Q9WLL2_HPBVO	Q9wll2 hepatitis b
38	1202	94.1	400	2 Q72030_HPBVO	Q72030 hepatitis b
39	1202	94.1	400	2 Q80GY1_HPBVO	Q80gy1 hepatitis b
40	1202	94.1	400	2 Q80MR0_HPBVO	Q80mr0 hepatitis b
41	1202	94.1	400	2 Q91GW9_HPBVO	Q91gw9 hepatitis b
42	1202	94.1	400	2 Q91GX1_HPBVO	Q91gx1 hepatitis b
43	1202	94.1	400	2 Q4FDF8_HPBVO	Q4fd8 hepatitis b
44	1202	94.1	400	2 Q4FDF8_HPBVO	Q4fd8 hepatitis b
45	1201	94.0	226	2 Q5G8L9_HPBVO	Q5g8l9 hepatitis b

ALIGNMENTS

RESULT 1

Q9DUJ4_HPBVO

ID Q9DUJ4_HPBVO PRELIMINARY; PRT; 226 AA.

AC Q9DUJ4; 16, Created

DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)

DE Small surface polypeptide.

GN Name=OG14;

OS Hepatitis B virus.

OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

OC Orthohepadnavirus.

OX NCBI_TaxID=10407;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Ogura Y., Kurosaki M., Asahina Y., Enomoto N., Marumo F., Sato C.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB029963; BAB20346.1; -; Genomic_DNA.

DR PIR; JQ2094; JQ2094.

DR PIR; JQ2095; JQ2095.

DR PIR; JQ2096; JQ2096.

DR PIR; JQ2097; JQ2097.

DR PIR; JQ2098; JQ2098.

DR PIR; JQ2099; JQ2099.

DR PIR; JQ2100; JQ2100.

DR PIR; JQ2101; JQ2101.

DR PIR; JQ2102; JQ2102.

DR PIR; JQ2106; JQ2106.

DR PIR; JQ2108; JQ2108.

DR PIR; JQ2109; JQ2109.

DR PIR; JQ2111; JQ2111.

DR PIR; JQ2112; JQ2112.

DR GO; GO:0016032; P: viral life cycle; IEA.

DR InterPro; IPR000349; Hepvir_surfa.

DR PANTHER; PTHR10832; Hepvir_surfa; 1.

DR Pfam; PF00695; VMSA; 1.

KW Antigen; Polyprotein.

SQ SEQUENCE 226 AA; 25375 MW; 7704565C4092E3BC CRC64;

Alignment Scores:			
Pred. No.:	9,276-100	Length:	226
Score:	1220.00	Matches:	222
Percent Similarity:	98.24%	Conservative:	1
Best Local Similarity:	97.80%	Mismatches:	3
Query Match:	95.46%	Indels:	2
DB:	2	Gaps:	0
US-10-761-006A-1_COPy_155_835 (1-681) x Q9DUJ4_HPBV0 (1-226)			
Qy	1	ATGCGAACAACAATCAGGATTCCTAGACCCCTGCTCGTGTATACAGCGGGTTC	60
Db	1	MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	20
Qy	61	TGTTGACAAGAATCCTCACATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	21	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn	40
Qy	120	TTTCTAGGGGAGACCCACGCTGTTCTCGGCCAAAAATTCGCAGTCCCAACCTCCAATCA	179
Db	41	PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHl	60
Qy	180	CTCACCAACCTCTTGCTCCAAATTTGCTCGGCTATCGCTGGATGTCTGGCGGCTTT	239
Db	60	sSerProThrSerCysProProIleCysProGlyTyTrArgTrpMetCysLeuArgArgPh	80
Qy	240	TATCATATTCCTTCATCCCTGCTGCTATGCTCATCTTCTTGTTGGTTCTTCTCGACTA	299
Db	80	eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy	100
Qy	300	CCAAGGTATGTTGCCGGTTTGCTCTACTTCCAGGAACATCAACACAGCAGCGGGCC	359
Db	100	rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr	120
Qy	360	ATGCAAGACTGCAGACTCTGTCTCAAGGAAACTTAAGTTCCCTCTTGTCTGTACT	419
Db	120	oCysIysThrCysThrSerProAlaGlnGlyAsnSerThrPheProSerCysCysCsth	140
Qy	420	AAACCTCTCGGACGGAACTGCATCTGTATTCCTCCATCCCATCTCTGGGCTTCCGACG	479
Db	140	rIlysProSerAspGlyAsnCystrCysIleProIleProSerSerTrpAlaPheAlaAr	160
Qy	480	ATTCTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT	539
Db	160	gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa	180
Qy	540	TCAGTGGTTCGTAGGGCTTCCCCACATGTTTGGCTTCAGTTATATATGATATGGTA	599
Db	180	lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy	200
Qy	600	TTGGGGCGCAAGTCTGTACAAATCTCTCAGTCCCTTTTACCTCTATTACCAATTTCTT	659
Db	200	rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh	220
Qy	660	TGTCCTTTGGGTATACATT	678
Db	220	eCysLeuTrpValTrile	226

RESULT 2

Q7TDS6	HBPV0
Q7TDS6	PRELIMINARY;
ID	Q7TDS6
AC	PRT; 400 AA.
AD	Q7TDS6;
DT	01-OCT-2003 (TREMBLrel. 25, Created)
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	PreS1/pres2 surface protein.
OS	Hepatitis B virus.
OC	Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC	Orthohepadnavirus.
OX	NCBI_TaxID=10407;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;

```

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167091; AA041300.1; -; Genomic DNA.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfAg.
DR PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR Pfam; PF00695; VMSA; 1.
DR KJ Antigen.
DR KX Antigen.
SQ SEQUENCE 400 AA; 43790 MW; 5D2FFC9D91CFE4D7 CRC64;

Alignment Scores:
Pred. No.: 4.23e-99 Length: 400
Score: 1213.00 Matches: 221
Percent Similarity: 97.36% Conservative: 0
Best Local Similarity: 97.36% Mismatches: 5
Query Match: 94.91% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1 COPY 155 835 (1-681) x 07TDS6 HPBV0 (1-400)

```

US-10-761-006A-1 COPY 155 835 (1-681) x O7TDS6 HPBV0 (1-400)

Qy	1	ATGAGAACACAA	CATCAGGATTC	TAGGACCCCT	GCCTGCTC	GTGCTTAC	AGCGGGGTTT	CTTC	60			
Db	175	MetGluAnThr	ThrSerGlyPhe	LeuGlyPro	LeuValLeu	GlnAla	GlyPhe	Phe	194			
Qy	61	TCGTTGACAAGA	ATCCTCACA	CAATACCG	CAGAGTCT	AGACTC	-TGGTGG	ACTTCTCTCAAT	119			
Db	195	LeuLeuThrArg	GlleLeu	ThrIlePro	GlnSerLeu	AspSer	TrpTrp	ThrSerLeu	Asn	214		
Qy	120	TTTCTAGGGGAG	CACCACTG	TGTTCTCT	GGCCAAAT	TCCGAGT	CCCCAAC	CTCCATCA	179			
Db	215	PheLeuGlyGly	AlaProThr	Cys-Pro	GlyGlnAsn	SerGln	SerPro	ThrSerAsn	Hi	234		
Qy	180	CTCACCAACNC	TTGTCTCC	CAATTTGT	CTCGGCTAT	CGCTGG	AGTGCT	CGCGGCTTT	239			
Db	234	sSerPro	ThrSerCys	ProProl	ileCys	ProGly	tyrArg	TrpMetCys	LeuUarg	Ph	254	
Qy	240	TATCATATTC	CTCTTCAT	CTCTGCT	AGCTCAT	CTCTTGT	TGGTTCT	CTTCGGACTA	299			
Db	254	eilePheLeu	PheIleLeu	LeuLeu	CysLeu	IlePheLeu	ValLeu	LeuLeu	Asp	Ty	274	
Qy	300	CCAGGTATGTT	GGCCGTTT	CTCTCTACT	TTCAGGAC	ATCAAC	CACAG	CACGGGGC	359			
Db	274	rgInGlyMet	LeuProVal	CysPro	LeuLeuPro	GlyThr	SerThr	ThrSer	ThrGly	Pr	294	
Qy	360	ATGCAAGAC	CTGCACGACT	CTCGTCT	CAAGAA	ACTCTAC	GTTCCTCT	TGTTGCTGTAC	419			
Db	294	oCysLys	ThrCys	ThrIlePro	AlaGln	GlyThr	SerMet	PhePro	SerCys	Cys	Th	314
Qy	420	AAACCTTCGAC	CGGAAC	TGCAC	TGTGAT	TCCCAT	CCCATCAT	CTCGGGCTTC	CGCAAG	479		
Db	314	rLysPro	SerAspGly	AsnCys	ThrCys	IlePro	llePro	SerSerTrp	AlaPhe	Ala	Ar	334
Qy	480	ATTCTATGGAG	TGGGGCTC	AGTCCG	TTCTCT	CTGGCTC	AGTTACT	AGTGCATTTGT	539			
Db	334	gPheLeuTrp	GluTrpAla	SerValArg	SerSerTrp	LeuSerLeu	LeuVal	Prophe	Val	354		
Qy	540	TCAGTGGTTC	TAGGGCTTT	CCCCAC	TGTTGG	CTTTCAG	TTATATG	ATGATGGGTA	599			
Db	354	lGlnTrpPhe	ValGlyLeu	SerPro	ThrValTrp	LeuSerVal	IleTrp	MetMet	Trp	Trp	374	
Qy	600	TTGGGGGGA	GTCTGTACA	CATCTTG	AGTCCCTTTT	TACCTCT	TATTA	CCAAATTTCTT	659			
Db	374	rTrpGlyArg	SerLeu	TyrAsn	IleLeu	SerPro	PheLeu	ProLeuLeu	Pro	IlePhe	Ph	394
Qy	660	TTGCTTTGG	TATACATT	678								
Db	394	eCysLeu	TrpVal	TyrIle	400							

RESULT 3

Q9DHA2 HPBV0
ID Q9DHA2 HPBV0 PRELIMINARY; PRT: 226 AA.

AC	Q9DHA2;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)

QY 1 ATGAGAACACAAATCCTCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGTTTTTC 60
 DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
 QY 61 TCGTTGACAGAAATCCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 DB 195 LeuLeuThrArgLeuThrLeuProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
 QY 120 TTTCTAGGGGAGACCCAGTGTCTGCGCCAAAATTCGACATCCCAACCTCCCAATCA 179
 DB 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
 QY 180 CTACCAACCTCTGCTCCCAATTTGCTCTCTACTTCCAGGAACATCAACCAACGACGGGCC 239
 DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgGph 254
 QY 240 TATCATATTCCTCTTCACTGCTGCTATGCCTCATCTCTGTTGTTCTCTGGACTA 299
 DB 254 eillePheLeuPheLeuLeuLeuCysLeuLeuPheLeuLeuValLeuLeuAspTy 274
 QY 300 CCAAGGTATGTTGCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 359
 DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
 QY 360 ATGCAAGACCTGACGACTCTGCTCAAGAACTCTACGTTTCCCTCTCTCTCTCTCTCT 419
 DB 294 oCysLeuThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
 QY 420 AAAACCTTCGAGGAGGAGTGCCTCTATTTCCATCCCACTATCTCTGGGCTTTCGCAAG 479
 DB 314 rLysProSerAspGlyAsnCysThrCysilleProIleProSerSerTrpAlaPheAlaAr 334
 QY 480 ATTCTATGAGGAGTGGGCTCAGTCCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
 DB 334 gTyrLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
 QY 540 TCAGTGTGCTAGGGCTTCCCACTGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 599
 DB 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
 QY 600 TTGGGGCGAGTCTGTACACATCTTGAGTCCCTTTTACCTTATACCAATTTCTTCT 659
 DB 374 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
 QY 660 TTGCTTTGGGTATACATT 678
 DB 394 eCysLeuTrpValTyrlle 400

RESULT 7

Q5SDK8 HPBV0
 ID Q5SDK8 HPBV0 PRELIMINARY; PRT; 400 AA.
 AC Q5SDK8;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DE Envelope protein.
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OK NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=He100;
 RA Song B.-C., Kim H., Kim S.-H., Cha C.-Y., Kook Y.-H., Kim B.-J.;
 RT "Comparison of full length sequences of hepatitis B virus isolates in
 RT hepatocellular carcinoma patients and asymptomatic carriers of
 RT Korea.";
 RL J. Med. Virol. 75:13-19(2005).
 DR EMBL; AY641560; AAV52022.1; -; Genomic DNA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0016032; P:viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfAg.

DR PANTHER; PTHR10832; Hepvir_surfAg; 1.
 DR Pfam; PF00695; yMSA; 1
 KW Antigen; Envelope protein.
 SQ SEQUENCE 400 AA; 43712 MW; 5C02A7936484BAC1 CRC64;

Alignment Scores:

Pred. No.: 1.18e-98 Length: 400
 Score: 1208.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.52% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q5SDK8_HPVB0 (1-400)

QY 1 ATGAGAACACAAATCCTCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGTTTTTC 60
 DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
 QY 61 TCGTTGACAGAAATCCTCACAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 DB 195 LeuLeuThrArgLeuThrLeuProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
 QY 120 TTTCTAGGGGAGACCCAGTGTCTGCGCCAAAATTCGACATCCCAACCTCCCAATCA 179
 DB 215 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
 QY 180 CTCACAACTCTGTTGCTCTCAATTTGCTGCTGCTGATCGTGTGATGTCTGGGGCTTT 239
 DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgGph 254
 QY 240 TATCATATTCCTCTTCACTGCTGCTATGCCTCATCTCTGTTGTTCTCTGGACTA 299
 DB 254 eillePheLeuPheLeuLeuLeuCysLeuLeuPheLeuLeuValLeuLeuAspTy 274
 QY 300 CCAAGGTATGTTGCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 359
 DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
 QY 360 ATGCAAGACCTGACGACTCTGCTCAAGAACTCTACGTTTCCCTCTCTCTCTCTCTCTCT 419
 DB 294 oCysLeuThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
 QY 420 AAAACCTTCGAGGAGGAGTGCCTCTATTTCCATCCCACTATCTCTGGGCTTTCGCAAG 479
 DB 314 rLysProSerAspGlyAsnCysThrCysilleProIleProSerSerTrpAlaPheAlaAr 334
 QY 480 ATTCTATGAGGAGTGGGCTCAGTCCGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
 DB 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
 QY 540 TCAGTGTGCTAGGGCTTCCCACTGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 599
 DB 354 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
 QY 600 TTGGGGCGAGTCTGTACACATCTTGAGTCCCTTTTACCTTATACCAATTTCTTCT 659
 DB 374 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
 QY 660 TTGCTTTGGGTATACATT 678
 DB 394 eCysLeuTrpValTyrlle 400

RESULT 8

Q91GW8 HPBV0
 ID Q91GW8 HPBV0 PRELIMINARY; PRT; 400 AA.
 AC Q91GW8;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Envelope protein.
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;

```
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN NUCLEOTIDE SEQUENCE.
RP Shi S., Dong J., Huangfu J.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY040799; AAK94658.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43682 MW; 283B5065EE6157CAF CRC64;

Alignment Scores:
Pred. No.: 1,18e-98 Length: 400
Score: 1208.00 Matches: 220
Percent Similarity: 97.36% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 5
Query Match: 94.52% Indels: 2
DB: Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x Q9PXL3_HPBVO (1-400)

QY 1 ATGGAGAACACATCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTC 60
DB 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194

QY 61 TCGTTCACAGAAATCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
DB 195 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214

QY 120 TTTCTAGGGGAGACCCACGTTTCTCGCCAAATTCGAGTCCCAACCTCCCAATCA 179
DB 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234

QY 180 CTCACCAACTCTTGTCCTCAATTTGTCCTGCTATCGCTGATGTGTCTGGGGGTTT 239
DB 234 sSerProThrSerCysProProIleCysProGlyTyArgTrpMetCysLeuArgArgPh 254

QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTGCTCTCTTCTGTTCTGTTCTGACTA 299
DB 254 elleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 274

QY 300 CCAAGTATGTTGCCGTTTGTCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
DB 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294

QY 360 ATGCAGACCTGCAGACTCTGCTCAAGAACTCTAGTTCCCTCTGTTCTGTTCTGTTAC 419
DB 294 oCysIleThrCysThrSerProAlaGlnGlyThrSerMetPheProSerCysCysTh 314

QY 420 AAAACCTTCGGACGGAACTGCACTTGTATTCCTCCATCCATCTCTGGGCTTTCGCAAG 479
DB 314 rIlySProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTATGGAGTGGGCGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTCCATTTGT 539
DB 480 ATTCTATGGAGTGGGCGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTCCATTTGT 539

Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
QY 540 TCAGTGGTTCGTAGGGCTTTCCCCACACTGTTTGGCTTTTTCAGTTATATGATGATGTTA 599
DB 354 IGLnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGCGCAAGTCTGTCAACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
DB 374 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 394
QY 660 TTGCTTTGGGTATACATT 678
DB 394 eCysLeuTrpValTyIle 400

RESULT 9
Q9PXL3_HPBVO
ID Q9PXL3_HPBVO PRELIMINARY; PRT; 226 AA.
AC Q9PXL3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Surface antigen.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=genotype C;
RX MEDLINE=20528425; PubMed=11074476;
RX DOI=10.1002/1096-9071(200012)62:4<471::AID-JMV12>3.3.CO;2-O;
RA Rokuhara A., Tanaka E., Yagi S., Mizokami M., Hashikura Y.,
RA Kawasaki S., Kiyosawa K.;
RT "De novo infection of hepatitis B virus in patients with orthotopic
RT liver transplantation: analysis by determining complete sequence of
RT the genome.";
RL J. Med. Virol. 62:471-478(2000).
DR EMBL; AB030517; BAA82654.1; -; Genomic DNA.
DR EMBL; AB030516; BAA82653.1; -; Genomic_DNA.
DR PIR; JQ2094; JQ2094.
DR PIR; JQ2095; JQ2095.
DR PIR; JQ2096; JQ2096.
DR PIR; JQ2097; JQ2097.
DR PIR; JQ2098; JQ2098.
DR PIR; JQ2099; JQ2099.
DR PIR; JQ2100; JQ2100.
DR PIR; JQ2101; JQ2101.
DR PIR; JQ2102; JQ2102.
DR PIR; JQ2106; JQ2106.
DR PIR; JQ2108; JQ2108.
DR PIR; JQ2109; JQ2109.
DR PIR; JQ2111; JQ2111.
DR PIR; JQ2112; JQ2112.
DR PIR; JQ2116; JQ2116.
DR GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 226 AA; 25394 MW; 835E71D8909A9C21 CRC64;

Alignment Scores:
Pred. No.: 1,33e-98 Length: 226
Score: 1207.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.44% Indels: 2
DB: Gaps: 0

US-10-761-006a-1_copy_155_835 (1-681) x Q9PXL3_HPBVO (1-226)

QY 1 ATGGAGAACACATCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTC 60
DB 1 ATGGAGAACACATCAGATTCCTAGGACCCCTGCTGTTTACAGGGGGGTTTTC 60
```


Db 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCCTTGACAGAAATCTCAATACCGCAGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
 Db 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGAGCACCACCGTGTCTCGSCAAATTCGCAGTCCCAACCTCCCAATCA 179
 Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH 60
 QY 180 CTCACCAACCTCTGTCTCCCAATTTGTCTGGCTATCGTGTGATGTGTGGGGGTTT 239
 Db 60 sSerProThrSerCysProProThrCysProGlyTyArgTrpMetCysLeuArgPh 80
 QY 240 TATCATATTTCTTTCATCTGTCTGTCTATGCTATGCTCTCTCTCTCTCTCTCTCT 299
 Db 80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTTCGGTGTTCCTCTACTTCCAGGAACATCAACACGACGACGGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
 QY 360 ATCAAGACCTGCACGACTCTCTCAAGGAAACTCTAGCTTTCCCTCTTGTGTGTATC 419
 Db 120 oCysLysThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGACGGAACCTGCACTGTATTCATCCATCCATCATCTGGGCTTCGCAAG 479
 Db 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGT 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 200
 QY 600 TTGGGGGGAAGTCTGTACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
 Db 200 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhe 220
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 220 eCysLeuTrpValTyIle 226

RESULT 10
 Q9QC26 HPBVR
 ID Q9QC26 HPBVR PRELIMINARY; PRT; 392 AA.
 AC Q9QC26;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Large envelope protein.
 OS Hepatitis B virus (subtype adr).
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 OX NCBI_TaxID=106820;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Dong J., Cheng J., Wang Q.H., Liu Y.Z., Xia X.B., Song H.B., Wang G.,
 RA Zhong Y.W., Shi S.S., Yang J.Z., Liu Y.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF191303; AAF05722.2; -; Genomic_DNA.
 DR PIR; JQ2094;
 DR PIR; JQ2095;
 DR PIR; JQ2096;
 DR PIR; JQ2097;
 DR PIR; JQ2098;
 DR PIR; JQ2099;
 DR PIR; JQ2100;
 DR PIR; JQ2101;
 DR PIR; JQ2102;

DR PIR; JQ2106; JQ2106.
 DR PIR; JQ2108; JQ2108.
 DR PIR; JQ2109; JQ2109.
 DR PIR; JQ2111; JQ2111.
 DR PIR; JQ2112; JQ2112.
 DR PIR; JQ2116; JQ2116.
 DR GO; GO:0015031; C:viral envelope; IEA.
 DR GO; GO:0015032; P:viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfac.
 DR PANTHER; PTHR10832; Hepvir_surfac; 1.
 DR Pfam; PF00695; vmsa; 1.
 KW Antigen; Envelope protein.
 SQ SEQUENCE 392 AA; 42699 MW; 0394CD9EFCCA1757 CRC64;

Alignment Scores:
 Pred. No.: 1,44e-98 Length: 392
 Score: 1207.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.44% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_155_835 (1-681) x Q9QC26 HPBVR (1-392)

QY 1 ATGGAGAACACACATCAGGATTCCTAGGACCCCTGCTGCTGTACAGCGGGGTTTTTC 60
 Db 167 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 186
 QY 61 TCGTTGACAGAAATCTCAATACCGCAGAGCTAGACTC-TGGTGGAGCTTCTCTCAAT 119
 Db 187 SerLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 206
 QY 120 TTTCTAGGGGAGCACCACCGTGTTCCTGCGCAAAATTCGCAGTCCCAACCTCCCAATCA 179
 Db 207 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnH 226
 QY 180 CTCACCAACCTCTGTCTCCCAATTTGTCTGGCTATGCTGCTGCTGCTGCTGCTGCT 239
 Db 226 sSerProThrSerCysProProThrCysProGlyTyArgTrpMetCysLeuArgPh 246
 QY 240 TATCATATTTCTTTCATCTGTCTGTCTATGCTGTCTCTCTCTCTCTCTCTCTCTCT 299
 Db 246 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 266
 QY 300 CCAAGGTATGTTCGGTGTTCCTCTACTTCCAGGAACATCAACACGACGACGGGCC 359
 Db 266 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 286
 QY 360 ATCAAGACCTGCACGACTCTCTCAAGGAAACTCTAGCTTTCCCTCTTGTGTGTATC 419
 Db 286 oCysLysThrCysThrThrProAlaGlnGlyThrSerMetPheProSerCysCysTh 306
 QY 420 AAAACCTTCGACGGAACCTGCACTGTATTCATCCATCCATCATCTGGGCTTCGCAAG 479
 Db 306 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 326
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 326 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 346
 QY 540 TCAGT 599
 Db 346 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrp 366
 QY 600 TTGGGGGGAAGTCTGTACAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
 Db 366 rTrpGlyProSerLeuTyAsnIleLeuSerProPheLeuProLeuLeuProIlePhe 386
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 386 eCysLeuTrpValTyIle 392

RESULT 11


```

Db      220 eCysLeuTrpValTyrIle 226
|||||
Db      80 ellellePheLeuPhelelleLeuLeuCysleullePheleulleValleuLeuAspTy 100
|||||
QY      300 CCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGAACATCAACACAGACGCGGCC 359
|||||
Db      100 rGInglyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
|||||
QY      360 ATGCAGACCTGCAGACTCCCTGCTCAAGAACTCTAGTTTCCCTCTTGTGCTGTAC 419
|||||
Db      120 oCyslysthrCysThrIleProAlaGInglyThrSerMetPheProSerCysCysTh 140
|||||
QY      420 AAAACCTTCGACGCGAACTGCATCTGTATTCCTCCATCCCATCATCTCGGCTTCCGCAAG 479
|||||
Db      140 rlysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAla 160
|||||
QY      480 ATTCTTATCGGAGTGGGCTCAGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
|||||
Db      160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
|||||
QY      540 TCAGTGGTTCGTAGGGCTTTCCTCCCACTCTTGGCTTTCAGTTATATGATGATGATGTA 599
|||||
Db      180 lGInTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
|||||
QY      600 TTGGGGCGCAAGTCTGTCAACATCTTGTAGTCCCTTTTACCTCTTATTAACCAATTTCTT 659
|||||
Db      200 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220
|||||
QY      660 TTGCTCTTGGGTATACATT 678
|||||
Db      220 eCysLeuTrpValTyrIle 226
|||||
RESULT 13
VMSA HPBVR STANDARD; PRT; 399 AA.
AC P03140;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Major surface antigen precursor.
GN Name=S;
OS Hepatitis B virus (subtype adr).
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=106820;
RN [1]
RX MEDLINE=8316819; PubMed=6300776;
RA Ono Y., Onda H., Sasada R., Igarashi K., Sugino Y., Nishioka K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adr.";
RL Nucleic Acids Res. 11:1747-1757(1983).
RN [2]
RX NUCLEOTIDE SEQUENCE OF 175-349.
RA MEDLINE=85107103; PubMed=3968537;
RA Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;
RT "Structural analysis of the gene coding for hepatitis B virus surface
antigen and its product.";
RL J. Gen. Virol. 66:195-200(1985).
CC -1- PTM: Myristoylation is essential for viral infectivity (By
similarity).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; V00867; CAA24234.1; ALT_INIT; Genomic_DNA.
DR PIR; A03705; SAVLA.
DR InterPro; IPR000349; Hepvir_surfag.
DR PANTHER; PTHR10832; Hepvir_surfag; 1.
DR Pfam; PF00695; VMSA; 1.
DR Antigen; Glycoprotein; Lipoprotein; Myristate.
KW

```

FT INIT_MET 0 0 By similarity.
 FT PROPEP 1 173 Major surface antigen.
 FT CHAIN 174 399 N-myristoyl glycine (by host) (By
 FT LIPID 1 1 similarity)
 FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 122 122 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 176 176 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 319 319 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 399 AA; 43554 MW; 3C7437A6EBDBD4E6 CRC64;

Alignment Scores:
 Pred. No.: 1.78e-98 Length: 399
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x VM5A_HPBVR (1-399)

QY 1 ATGGAGAACACAAATCAGAGTTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTTC 60
 Db 174 MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 193
 QY 61 TCGTTGACAAGAATCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 194 LeuLeuThrArgileLeuThrileProGlnSerLeuAspSerTrpTrpThrSerLeuAen 213
 QY 120 TTTCTAGGGGAGCACCACGCTGTTCTCGGCCAAATTCGAGTCCCGAACCTCCATCA 179
 Db 214 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 233
 QY 180 CTCACCAACTCTGTCCTCAATTTGCTCGCTATCGCTGATGCTGCTGCTGCTGCTGCT 239
 Db 233 sSerProThrSerCysProProleCysProGlyTyArgTrpMetCysLeuArgPh 253
 QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTCTTCTTGTGGTTCTCTGACTA 299
 Db 253 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuValLeuLeuAspTy 273
 QY 300 CCAAGTATGTGCGCGTGTGCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
 Db 273 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 293
 QY 360 ATGCAAGACTGCGACACTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
 Db 293 oCyslysthrCysThrileProAlaGlnGlyThrSerMetPheProSerCysCysTh 313
 QY 420 AAAACCTTCGGACGGAACTGCACCTTGTATTCCTCCATCCCATCATCTCGGCTTTCGCAAG 479
 Db 313 rLysProSerAspGlyAsnCysThrCysileProileProSerSerTrpAlaPheAlaAr 333
 QY 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT 539
 Db 333 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 353
 QY 540 TCAGTGGTTCGTAGGGCTTTCCCGCACTGTTGGCTTTCCAGTTATATGATGATGTGTA 599
 Db 353 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 373
 QY 600 TTGGGGCGGAAGTCTGTACAACATCTTGAGTCCCTTTTTTACCTCTATTATCAATTTCTT 659
 Db 373 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuLeuProilePhePh 393
 QY 660 TTGTCTTTGGGTATACATT 678
 Db 393 eCysLeuTrpValTyIle 399

RESULT 14

Q68RP7_HPBVO

ID Q68RP7_HPBVO PRELIMINARY; PRT; 400 AA.

AC Q68RP7;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Pres1/pres2 surface.
 OS Hepatitis B virus.
 OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
 OC Orthohepadnavirus.
 RN NCBI_TaxID=10407;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Liu C.-J.; Chen P.-J.; Chen D.-S.; Kao J.-H.; Lai M.-Y.;
 RT "Origin of Serum Hepatitis B Virus in Acute Exacerbation: Comparison
 RT with HBV in the liver and from other Exacerbation.";
 RL Hepatology 40:310-317(2004).
 DR EMBL; AY596108; AAU01945.1; -; Genomic DNA.
 DR GO; GO:0016032; P:Viral life cycle; IEA.
 DR InterPro; IPR000349; Hepvir_surfaG.
 DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
 DR Pfam; PF00695; VM5A; 1.
 KW Antigen.
 SQ SEQUENCE 400 AA; 43768 MW; DBDC9D063A569678 CRC64;

Alignment Scores:
 Pred. No.: 1.78e-98 Length: 400
 Score: 1206.00 Matches: 220
 Percent Similarity: 96.92% Conservative: 0
 Best Local Similarity: 96.92% Mismatches: 6
 Query Match: 94.37% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q68RP7_HPBVO (1-400)

QY 1 ATGGAGAACACAAATCAGAGTTCCTAGGACCCCTCGTGTGTACAGCGGGGTTTTTC 60
 Db 175 MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 194
 QY 61 TCGTTGACAAGAATCTCAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 195 LeuLeuThrArgileLeuThrileProGlnSerLeuAspSerTrpTrpThrSerLeuAen 214
 QY 120 TTTCTAGGGGAGCACCACGCTGTTCTCGGCCAAATTCGAGTCCCGAACCTCCATCA 179
 Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAenSerGlnSerProThrSerAsnHi 234
 QY 180 CTCACCAACTCTGCTCTCAATTTGCTCGGCTATCGCTGATGCTGCTGCTGCTGCT 239
 Db 234 sSerProThrSerCysProProleCysProGlyTyArgTrpMetCysLeuArgPh 254
 QY 240 TATCATATTCTCTTCATCTGCTGCTATGCTCATCTTCTTGTGGTTCTCTGACTA 299
 Db 254 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuValLeuLeuAspTy 274
 QY 300 CCAAGTATGTGCGCGTGTGCTCTACTTCCAGGAACATCAACACGACGCGGGCC 359
 Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
 QY 360 ATGCAAGACTGCGACACTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
 Db 294 oCyslysthrCysThrileProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
 QY 420 AAAACCTTCGGACGGAACTGCACCTTGTATTCCTCCATCCCATCATCTCGGCTTTCGCAAG 479
 Db 314 rLysProSerAspGlyAsnCysThrCysileProileProSerSerTrpAlaPheAlaAr 334
 QY 480 ATTCCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT 539
 Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 354
 QY 540 TCAGTGGTTCGTAGGGCTTTCCCGCACTGTTGGCTTTCCAGTTATATGATGATGTGTA 599
 Db 354 iGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
 QY 600 TTGGGGCGGAAGTCTGTACAACATCTTGAGTCCCTTTTTTACCTCTATTATCAATTTCTT 659

Db 374 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePhe 394
QY 660 TTGTCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyrIle 400

RESULT 15

Q7TDR4_HPBVO
ID Q7TDR4_HPBVO PRELIMINARY; PRT; 400 AA.
AC Q7TDR4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PresI/pres2 surface protein.
OS Hepatitis B virus.
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Liu C.-J., Chen P.-J., Lai M.-Y., Kao J.-H., Chen D.-S.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY167095; AA041312.1; -; Genomic DNA.
DR GO; GO:0016032; P: viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfaG.
DR PANTHER; PTHR10832; Hepvir_surfaG; 1.
DR Pfam; PF00695; vmsA; 1.
KW Antigen.
SQ SEQUENCE 400 AA; 43694 MW; CCF173EA4365E414 CRC64;

Alignment Scores:

Pred. No.: 1.78e-98 Length: 400
Score: 1206.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x Q7TDR4_HPBVO (1-400)

QY 1 ATGAGACACAAATCAGGATTCCTAGGACCTCTGCTGTACAGCGGGGTTTTC 60
Db 175 MetGluAsnThrThrSerGlyPheLeuGlyProLeuValLeuGlnAlaGlyPhePhe 194
QY 61 TCGTTGACAAATCCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 195 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 214
QY 120 TTCTAGGGGAGCACCACGCTGTTCTCGGCAAAATTCGAGTCCCAACCTCAATCA 179
Db 215 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 234
QY 180 CTCACCAACCTCTGCTCCAAATTTGCTCGGCTATCGCTGGATGTCGTGGCGGCTTT 239
Db 234 sSerProThrSerCysProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 254
QY 240 TATCATATTCCTCTCATCTGCTGCTATGCTCATCTTCTTGTGTTCTCTGGACTA 299
Db 254 eIleIlePheLeuPheIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 274
QY 300 CCAAGTATGTTGCCCGTTTCTCTCTACTTCCAGGAACATCAACACACGACGCGGCC 359
Db 274 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 294
QY 360 ATGCAAGACCTGCGACGACTCTGCTCAAGGAAACTCTACGTTTCCCTCTGCTGTGAC 419
Db 294 oCysAlaThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 314
QY 420 AARACCTTCGACGGAACCTGCTGATTCCTCATCCATCCATCTGGGCTTCGCAAG 479
Db 314 rLysProSerAspGlyAsnCysThrCysIlePheIleProSerSerTrpAlaPheAlaAr 334

QY 480 ATTCTCTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCCATTGT 539
Db 334 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 354
QY 540 TCAGTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAGTTATATGATGATGGA 599
Db 354 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 374
QY 600 TTGGGGGCGAGTCTGTACACATCTTGGTCCCTTTTACCTCTATTACCAATTTCCT 659
Db 374 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuPhePhe 394
QY 660 TTGTCTTTGGGTATACATT 678
Db 394 eCysLeuTrpValTyrIle 400

Search completed: December 27, 2005, 21:00:24

Job time : 234.454 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:42:07 ; Search time 38.59 Seconds
(without alignments)
3395.886 Million cell updates/sec

Title: US-10-761-006A-1_COPY_155_835

Perfect score: 1278
Sequence: 1 ATGGAGACACACATCAGG.....GTCTTTGGGTATACATTAA 681

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192901_15763/app_query.fasta_1.1102
-DB=PIR -OPMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10761006 @CNG 1 1 77 @runat_27122005_192901_15763 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1206	94.4	400	1 SAVLA	large surface anti
2	1201	94.0	226	2 JQ2112	surface antigen -
3	1201	94.0	226	2 JQ2106	surface antigen -
4	1200	93.9	226	2 JQ2110	surface antigen -
5	1198	93.7	226	2 JQ2108	surface antigen -
6	1197	93.7	226	1 JQ1576	major surface anti
7	1197	93.7	226	1 SAVLAD	major surface anti
8	1197	93.7	226	2 JQ2104	surface antigen -
9	1197	93.7	226	2 JQ2111	surface antigen -
10	1197	93.7	226	2 JQ2098	surface antigen -
11	1196	93.6	226	1 SAVLN1	major surface anti
12	1192	93.3	226	2 JQ2115	surface antigen -
13	1191	93.2	226	2 JQ2094	surface antigen -
14	1190	93.1	226	2 JQ2101	surface antigen -

15	1190	93.1	384	2 T13469	large surface anti
16	1189	93.0	226	2 JQ2103	surface antigen -
17	1189	93.0	226	2 JQ2114	surface antigen -
18	1188	93.0	226	1 JQ1581	major surface anti
19	1188	93.0	226	2 JQ2109	surface antigen -
20	1188	93.0	226	2 JQ2105	surface antigen -
21	1187	92.9	226	2 JQ2116	surface antigen -
22	1187	92.9	226	2 JQ2100	surface antigen -
23	1186	92.8	226	2 JQ2102	surface antigen -
24	1186	92.8	400	2 S35528	surface antigen -
25	1185	92.7	226	2 JQ2099	surface antigen -
26	1184	92.6	226	1 JQ1574	major surface anti
27	1184	92.6	226	1 SAVLAR	major surface anti
28	1182	92.5	226	2 JQ2095	surface antigen -
29	1180	92.3	226	2 JQ2097	surface antigen -
30	1174	91.9	226	2 JQ2054	surface antigen -
31	1171	91.6	226	2 JQ2052	surface antigen -
32	1168	91.4	226	1 SAVLHV	major surface anti
33	1167	91.3	226	2 JQ2050	surface antigen -
34	1166	91.2	226	2 JQ2096	surface antigen -
35	1163	91.0	226	1 JQ1577	major surface anti
36	1163	91.0	226	2 JQ2053	surface antigen -
37	1163	91.0	226	2 JQ2051	surface antigen -
38	1163	91.0	400	1 JQ1575	major surface anti
39	1163	91.0	400	1 SAVLVD	large surface anti
40	1162	90.9	226	2 JQ2048	surface antigen -
41	1158	90.6	226	2 JQ2113	surface antigen -
42	1157	90.5	226	2 JQ2055	surface antigen -
43	1156	90.5	389	1 SAVLVE	large surface anti
44	1149	89.9	226	2 JQ2078	surface antigen -
45	1149	89.9	226	2 JQ2045	surface antigen -

ALIGNMENTS

RESULT 1

SAVLA

large surface antigen - hepatitis B virus (subtype adr)
N;Contains: major surface antigen; middle surface antigen

C;Species: hepatitis B virus, HBV

C;Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text_change 09-Jul-2004

C;Accession: A03705; S04569; JQ2107; PQ0608

R;Ono, Y.; Onda, H.; Sabada, R.; Igarashi, K.; Sugino, Y.; Nishioka, K.

Nucleic Acids Res. 11, 1747-1757, 1983

A;Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype

A;Reference number: A93460; MUID:83168919; PMID:6300776

A;Accession: A03705

A;Molecule type: DNA

A;Residues: 1-400 <ON>

A;Cross-references: UNIPROT:P03140; UNIPARC:UPI00001389C1; GB:V00867

R;Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.

Nucleic Acids Res. 17, 2124, 1989

A;Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subty

A;Reference number: S04568; MUID:89183619; PMID:2928116

A;Accession: S04569

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-50, 'R', '52-66, 'YP', '69-129, 'V', '131-142, 'P', '144-164, 'S', '166-176, 'S', '178-338,

A;Cross-references: UNIPARC:UPI000017496C; EMBL:X14193

R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnius

J. Gen. Virol. 74, 1341-1348, 1993

A;Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A;Reference number: JQ2044; MUID:93329382; PMID:8336122

A;Accession: JQ2107

A;Molecule type: DNA

A;Residues: 175-400 <NOR>

A;Cross-references: UNIPARC:UPI0000033737

R;Norder, H.; Courouce, A.M.; Magnius, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A;Title: Molecular basis of hepatitis B virus serotype variations within the four major

A;Reference number: PQ0453; MUID:93107848; PMID:1469353

A;Accession: PQ0608

```

A;Molecule type: DNA
A;Residues: 275-354 <NO2>
A;Cross-references: UNIPARC:UPI000017496D
A;Experimental source: subtype adrq+, Bau
C;Genetics:
A;Gene: pre-S1/pre-S2/S
C;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen
F;120-400/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>
F;175-400/Product: major surface antigen (gene S) #status predicted <MSA>
F;15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2,83e-100 Length: 400
Score: 1205.00 Matches: 220
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 6
Query Match: 94.37% Indels: 2
DB: 1 Gaps: 0

```



```
|||||
Db      220 eCysLeuTrpValTyrIle 226

RESULT 3
JQ2106
surface antigen - hepatitis B virus (subtype adr, strain OLL)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2106
R:Norde, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup C
A:Accession: JQ2106
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q8UFP3; UNIPROT:Q91546; UNIPROT:Q9B6S4; UNIPROT:Q9DUK5; UNIP
PROT:Q9PXA0; UNIPROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99H08; UNIPROT:Q9B
DR80; UNIPROT:Q8VIM3; UNIPROT:Q9DTC4; UNIPROT:Q9DQK2; UNIPROT:Q9DUK6; UNIPROT:Q918J5; UN
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      8.1e-100      Length:      226
Score:          1201.00      Matches:      219
Percent Similarity: 96.48%      Conservative: 0
Best Local Similarity: 96.48%      Mismatches: 7
Query Match:      93.97%      Indels:      2
DB:              2          Gaps:      0

US-10-761-006a-1_COPY_155_835 (1-681) x JQ2106 (1-226)

QY      1  ATGGAGAACAACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTTTTC 60
Db      1  MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20

QY      61  TCGTTGACAGAATCCTCACATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db      21  TrpLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40

QY      120  TTTCTAGGGGGAGCACCACGCTGTTCTGCGCTATCGCTGGATGTCTCGCGGGTTT 179
Db      41  PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60

QY      180  CTCACCAACCTCTTGCTCCAAATTTGCTGCTGCTATCGCTGGATGTCTCGCGGGTTT 239
Db      60  sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80

QY      240  TATCATATTCCTTTCATCTGCTGCTATGCTCATCTCTTCTGTTGTTCTTCTGACTA 299
Db      80  eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100

QY      300  CCAAGTATGTCCTCCGTTGCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
Db      100  rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120

QY      360  ATGCAAGACTGCACGACTCCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
Db      120  oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140

QY      420  AAAACCTTCGGACGGAAACTGTATTCATCCATCCCATCATCTCTGGCTTCGCAAG 479
Db      140  rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160

QY      480  ATTCCTATGGAGTGGGCTCAGTCGGCTTCTCTGGCTCAGTTACTAGTGCATTGT 539
Db      160  gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 180

QY      540  TCAGTGTGTCGTAGGGCTTTCCCGCCACTGTTTGGCTTTTACGTTATATGATGATGGTA 599
Db      180  iGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
```

```
|||||
QY      600  TTGGGGCGAAGTCTGTACACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
Db      200  rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220

QY      660  TTGTCTTTGGGTATACATT 678
Db      220  eCysLeuTrpValTyrIle 226

RESULT 4
JQ2110
surface antigen - hepatitis B virus (subtype adr, strain pNDR260)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
C:Accession: JQ2110
R:Norde, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup C
A:Accession: JQ2110
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9QMK3; UNIPARC:UPI000017841E
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.:      9.96e-100      Length:      226
Score:          1200.00      Matches:      218
Percent Similarity: 96.92%      Conservative: 2
Best Local Similarity: 96.04%      Mismatches: 6
Query Match:      93.90%      Indels:      0
DB:              2          Gaps:      0

US-10-761-006a-1_COPY_155_835 (1-681) x JQ2110 (1-226)

QY      1  ATGGAGAACAACATCAGGATTCCTAGGACCCCTGCTCGTGTACAGGCGGGTTTTTC 60
Db      1  MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20

QY      61  TCGTTGACAGAATCCTCACATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db      21  LeuLeuThrLysIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40

QY      120  TTTCTAGGGGGAGCACCACGCTGTTCTGCGCAAAATTCGACGTCCCAACCTCCCAATCA 179
Db      41  PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60

QY      180  CTCACCAACCTCTTGCTCCAAATTTGCTGCTATCGCTGGATGTCTCGCGGGTTT 239
Db      60  sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80

QY      240  TATCATATTCCTTTCATCTGCTGCTATGCTCATCTCTTCTGTTGTTCTTCTGACTA 299
Db      80  eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuValLeuLeuAspTy 100

QY      300  CCAAGTATGTCCTCCGTTGCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
Db      100  rGlnGlyMetLeuProValCysProLeuLeuProGlyThrThrThrSerThrGlyPr 120

QY      360  ATGCAAGACTGCACGACTCCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
Db      120  oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140

QY      420  AAAACCTTCGGACGGAAACTGTATTCATCCATCCCATCATCTCTGGCTTCGCAAG 479
Db      140  rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160

QY      480  ATTCCTATGGAGTGGGCTCAGTCGGCTTCTCTGGCTCAGTTACTAGTGCATTGT 539
Db      160  gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValPropheVa 180
```


Db 60 sSerProThrSerCysProPheCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTCATCTCTTGTGTGGTCTCTCTGACTA 299
Db 80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTTGCCGGTTGCTCTACTTCCAGGAACATCAACACACGACGACGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
QY 360 ATGCAAGACCTGCACGACTCTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db 120 oCyslyThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGCTATGATTCCTCATCTCCATCATCTCGGCTTTCGAAG 479
Db 140 rlyProSerAspGlyAsnCysThrCysileProileProSerSerTrpAlaPheAlaAr 160
QY 480 ATTCTATGCGAGTGGGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGCT 539
Db 160 gPheLeuTrpGluTrpAlaSerAlaArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTGTTGCTAGGCTTTCCTCCACTGTTTGGCTTTCAGTTATATGATGATGGTA 599
Db 180 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTCTACAACTCTGATGCTCTCTTTTACCTCTATTAACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuProleilePhePh 220
QY 660 TTGCTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226
RESULT 7
SAVLAD
major surface antigen - hepatitis B virus (subtype ad)
C:Species: hepatitis B virus, HBV
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: PL0053
R:Okamoto, H.; Omi, S.; Wang, Y.; Itoh, Y.; Tauda, F.; Tanaka, T.; Akahane, Y.; Miyakawa
Mol. Immunol. 26, 197-205, 1989
A:Title: The loss of subtypic determinants in alleles, d/y or w/r, on hepatitis B surfac
A:Reference number: PL0053; MUID:89143494; PMID:2465492
A:Accession: PL0053
A:Molecule type: DNA
A:Residues: 1-226 <OKA>
A:Cross-references: UNIPROT:P31868; UNIPARC:UPI00001389BA; GB:M27765; NID:9329706; PIDN:
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
P:3/Binding site: carbohydrate (Aen) (covalent) #status predicted
Alignment Scores:
Pred. No.: 1.85e-99 Length: 226
Score: 1197.00 Matches: 218
Percent Similarity: 96.04% Conservative: 0
Best Local Similarity: 96.04% Mismatches: 8
Query Match: 93.66% Indels: 2
DB: 1 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x SAVLAD (1-226)
QY 1 ATGGAACACACATCAGATTCTAGGACCCCTCTGCTGTATACAGGGGGTTC 60
Db 1 MetGluAsnIleThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
QY 61 TCCTGACAGATCTCTCAATACCGACGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
Db 21 LeuLeuThrArgileLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40

QY 120 TTTCTAGGGGAGCACCACCGAGTGTCTCGGCCAAAATTCGACAGTCCCCAACTCCCAATCA 179
Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHl 60
QY 180 CTCACCAACTCTTGTCTCTCCAAATTTGCTCGCTATCGCTGATGTGTCTGGGGTTC 239
Db 60 sSerProThrSerCysProPheCysProGlyTyrArgTrpMetCysLeuArgArgPh 80
QY 240 TATCATATTCCTCTTCATCTGCTGCTATGCTCATCTCTTGTGTGTGTCTCTGACTA 299
Db 80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
QY 300 CCAAGGTATGTTGCCGGTTGCTCTACTTCCAGGAACATCAACACACGACGACGGGCC 359
Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrGlyPr 120
QY 360 ATGCAAGACCTGCACGACTCTCTGCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTGTAC 419
Db 120 oCyslyThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysCysTh 140
QY 420 AAAACCTTCGGACGGAACCTGCTATGATTCCTCATCTCCATCATCTCGGCTTTCGAAG 479
Db 140 rlyProSerAspGlyAsnCysThrCysileProileProSerSerTrpAlaPheAlaAs 160
QY 480 ATTCTATGCGAGTGGGCTCAGTCCGTTTCTCTCGCTCAGTTTACTAGTGCATTGCT 539
Db 160 nPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
QY 540 TCAGTGTGTTGCTAGGCTTTCCTCCACTGTTTGGCTTTCAGTTATATGATGATGGTA 599
Db 180 lGlnTrpPheAlaGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
QY 600 TTGGGGCGAAGTCTCTACAACTCTGATGCTCTCTTTTACCTCTATTAACCAATTTCTT 659
Db 200 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuProleilePhePh 220
QY 660 TTGCTCTTTGGGTATACATT 678
Db 220 eCysLeuTrpValTyIle 226
RESULT 8
JQ2104
surface antigen - hepatitis B virus (subtype adr, strain PHBV1-1)
C:Species: hepatitis B virus, HBV
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C:Accession: JQ2104
R:Norde, H.; Hammas, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus
J. Gen. Virol. 74, 1341-1348, 1993
A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origi
A:Reference number: JQ2044; MUID:93329382; PMID:8336122
A:Contents: genogroup C
A:Accession: JQ2104
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q81100; UNIPROT:Q930V8; UNIPARC:UPI00000F3B
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
Alignment Scores:
Pred. No.: 1.85e-99 Length: 226
Score: 1197.00 Matches: 218
Percent Similarity: 96.48% Conservative: 1
Best Local Similarity: 96.04% Mismatches: 7
Query Match: 93.66% Indels: 2
DB: 2 Gaps: 0
US-10-761-006a-1_COPY_155_835 (1-681) x JQ2104 (1-226)
QY 1 ATGGAACACACATCAGATTCTAGGACCCCTCTGCTGTATACAGGGGGTTC 60
Db 1 MetGluSerThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20

QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThraGilleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGAGACCCACGTTGTCCTGGCCAAATTCGAGTCCCACTCCAATCA 179
 Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTGTCCTCAATTTGTCCTGCGCTATCGCTGGATGCTGCGCGCGTTT 239
 Db 60 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuAargArgph 80
 QY 240 TATCATATTCCTCTTCATCTGCTCTATGCTCATCTCTTCTTTGGTTCCTTCGACTA 299
 Db 80 eileilePheLeuPheileLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTCGCGTTGTCCTACTTCCAGGAAACATCAACACGACGCGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGACGACTCCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
 Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysEth 140
 QY 420 AAAACCTTCGGACGGAACCTGACACTGTTGATTTCCCATCTCCATCTGCGGCTTCGCAAG 479
 Db 140 rlySerProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr 160
 QY 480 ATTCTATGGGAGTGGGCTCAGTCCGTTTCCTGCTGCTAGTTTACGTGACATGCTGTA 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 600 TTGGGGGGGAAAGTCTGTACAAATCTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
 Db 200 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuLeuProllePhePh 220

RESULT 9
 JQ2111

surface antigen - hepatitis B virus (subtype adr, strain Bes)
 C:Species: hepatitis B virus, HBV
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004
 C:Accession: JQ2111
 R:Norder: H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,
 J. Gen. Virol. 74, 1341-1348, 1993
 A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin
 A:Reference number: JQ2044; MUID:93329382; PMID:8336122
 A:Contents: genogroup C
 A:Accession: JQ2111
 A:Molecule type: DNA
 A:Residues: 1-226 <NOR>
 A:Cross-references: UNIPROT:Q91546; UNIPROT:Q9B6S4; UNIPROT:Q9DUK5; UNIPROT:Q992I5; UNIF
 PROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99H08; UNIPROT:Q9E950; UNIPROT:Q9D
 17V2; UNIPROT:Q9V1M3; UNIPROT:Q9DTC4; UNIPROT:Q9DKQ2; UNIPROT:Q918J5; UN
 C:Genetics:
 A:Gene: S
 C:Keywords: surface antigen

Alignment Scores:
 Pred. No.: 1.85e-99 Length: 226
 Score: 1197.00 Matches: 218
 Percent Similarity: 96.48% Conservative: 1
 Best Local Similarity: 96.04% Mismatches: 7
 Query Match: 93.66% Indels: 2
 DB: Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x JQ2111 (1-226)
 QY 1 ATGAGAACACAAATCAGGATTCCTAGGACCCCTGCTCGTGTACAGCGGGGTTTTTC 60
 Db 1 MetAspLysThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGTCTAGACTC-TGGTGGACTTCTCTCAAT 119
 Db 21 LeuLeuThraGilleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGAGACCCACGTTGTCCTGGCCAAATTCGAGTCCCACTCCAATCA 179
 Db 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTGTCCTCAATTTGTCCTGCGCTATCGCTGGATGCTGCGCGCGTTT 239
 Db 60 sSerProThrSerCysProProlleCysProGlyTyArgTrpMetCysLeuAargArgph 80
 QY 240 TATCATATTCCTCTTCATCTGCTCTATGCTCATCTCTTCTTTGGTTCCTTCGACTA 299
 Db 80 eileilePheLeuPheileLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTCGCGTTGTCCTACTTCCAGGAAACATCAACACGACGCGGCC 359
 Db 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGACGACTCCTGCTCAAGGAAACTCTACGTTTCCCTCTTGTGCTGTAC 419
 Db 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysEth 140
 QY 420 AAAACCTTCGGACGGAACCTGACACTGTTGATTTCCCATCTCCATCTGCGGCTTCGCAAG 479
 Db 140 rlySerProSerAspGlyAsnCysThrCysIleProlleProSerSerTrpAlaPheAlaAr 160
 QY 480 ATTCTATGGGAGTGGGCTCAGTCCGTTTCCTGCTGCTAGTTTACGTGACATGCTGTA 539
 Db 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuValProPheVa 180
 QY 540 TCAGTGGTTCGTAGGGCTTCCCACTGTTGGCTTTCAGTTATATGATGATGCTGTA 599
 Db 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
 QY 600 TTGGGGGGGAAAGTCTGTACAAATCTTGAGTCCCTTTTACCTCTATTACCAATTTTCTT 659
 Db 200 rTrpGlyProSerLeuTyAsnileLeuSerProPheLeuProLeuLeuProllePhePh 220

660 TTGTCCTTTGGGTATACATT 678

220 eCysLeuTrpValTyrlle 226

RESULT 10

JQ2098

surface antigen - hepatitis B virus (subtype adr, strain Bau)

C:Species: hepatitis B virus, HBV

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004

C:Accession: JQ2098

R:Norder: H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus

J. Gen. Virol. 74, 1341-1348, 1993

A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A:Reference number: JQ2044; MUID:93329382; PMID:8336122

A:Contents: genogroup C

A:Accession: JQ2098

A:Molecule type: DNA

A:Residues: 1-226 <NOR>

A:Cross-references: UNIPROT:Q8UYF3; UNIPROT:Q91546; UNIPROT:Q9B6S4; UNIPROT:Q9DUK5; UNIF

PROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99H08; UNIPROT:Q9E950; UNIPROT:Q9D

17V2; UNIPROT:Q9V1M3; UNIPROT:Q9DTC4; UNIPROT:Q8V1M3; UNIPROT:Q9DKQ2; UNIPROT:Q918J5; UN

C:Genetics:

A:Gene: S

C:Keywords: surface antigen

Alignment Scores:

Pred. No.: 1,85e-99 Length: 226
 Score: 1197.00 Matches: 218
 Percent Similarity: 96.48% Conservative: 1
 Best Local Similarity: 96.04% Mismatches: 7
 Query Match: 93.66% Indels: 2
 DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x JQ2098 (1-226)

QY 1 ATGGAGAACACAAATCCTAGGACCCCTGCTGCTGTTACAGGGGGGTTTTTC 60
 DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGGAGCACCACGTGTTCTGCGCCAAATTCGACGTCCCAACCTCCCAATCA 179
 DB 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnLeuGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTTGCTCCTCCAAATTTGCTGCTGCTATCGCTGGATGTCGTGGGGGCTTT 239
 DB 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgGPh 80
 QY 240 TATCATATTCCTTTCATCTGCTGCTATGCTCATCTTCTGTTGTTGTTCTCTGGAATA 299
 DB 80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTTGCCGCTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGCGACGACTCTGCTCAAGGAACTCTACGTTCCCTCTGTTGCTGTAC 419
 DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGGACGGAACTGCACTGTATTCATCCATCCCATCATCTCGGGCTTCGCAAG 479
 DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTCTCTGCTGCTCAGTTTACTAGTGCATTTGT 539
 DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
 QY 540 TCAGTGTTCGTAGGGCTTTCCCGACTGTTGGCTTTCAAGTTATATGGATGATGTTGTA 599
 DB 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
 QY 600 TTGGGGCGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTTATACCAATTTCTT 659
 DB 200 rTrpGlyProSerLeuTyraAsnIleLeuAsnProPheLeuProLeuLeuProIlePhePh 220
 QY 660 TTGCTTTGGGTATACATT 678
 DB 220 eCysLeuTrpValTyriIle 226

RESULT 11

SAVLN1

major surface antigen - hepatitis B virus (subtype adr, strain NC-1)

C;Species: hepatitis B virus, HBV

A;Note: host Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004

C;Accession: JCI1002

R;Qi, Z.H.; Yan, J.; Xiong, W.J.; Cai, L.W.

Chinese Biochem. J. 4, 201-209, 1998

A;Title: Determination of the nucleotide sequence and studies on the structure of hepatitis B virus surface antigen

A;Reference number: JCI1002

A;Molecule type: DNA

A;Residues: 1-226 <Q1Z>

A;Cross-references: UNIPROT:P30019; UNIPARC:UPI00001389BE

C;Genetics:

A;Gene: S
 C;Superfamily: hepatitis B virus surface antigen
 C;Keywords: surface antigen

Alignment Scores:
 Pred. No.: 2,28e-99 Length: 226
 Score: 1196.00 Matches: 218
 Percent Similarity: 96.04% Conservative: 0
 Best Local Similarity: 96.04% Mismatches: 8
 Query Match: 93.58% Indels: 2
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x SAVLN1 (1-226)

QY 1 ATGGAGAACACAAATCCTAGGACCCCTGCTGCTGTTACAGGGGGGTTTTTC 60
 DB 1 MetGluAsnThrAlaSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20
 QY 61 TCGTTGACAGAAATCCTCACAAATACCGCAGAGCTAGACTC-TGGTGGACTTCTCTCAAT 119
 DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAsn 40
 QY 120 TTTCTAGGGGGAGCACCACGTGTTCTGCGCCAAATTCGACGTCCCAACCTCCCAATCA 179
 DB 41 PheLeuGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60
 QY 180 CTCACCAACCTCTTGCTCCTCCAAATTTGCTGCTGCTATCGCTGGATGTCGTGGGGGCTTT 239
 DB 60 sSerProThrSerCysProIleCysProGlyTyArgTrpMetCysLeuArgGPh 80
 QY 240 TATCATATTCCTTTCATCTGCTGCTATGCTCATCTTCTGTTGTTGTTCTCTGGAATA 299
 DB 80 eIleIlePheLeuPheIleLeuLeuLeuCysLeuIlePheLeuLeuValLeuLeuAspTy 100
 QY 300 CCAAGGTATGTTGCCGCTTGTCTCTACTTCCAGGAACATCAACACACGACGCGGGCC 359
 DB 100 rHISGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120
 QY 360 ATGCAAGACCTGCGACGACTCTGCTCAAGGAACTCTACGTTCCCTCTGTTGCTGTAC 419
 DB 120 oCysLysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140
 QY 420 AAAACCTTCGGACGGAACTGCACTGTATTCATCCATCCCATCATCTCGGGCTTCGCAAG 479
 DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaAr 160
 QY 480 ATTCTATGGAGTGGGCTCAGTCCGTTCTCTGCTGCTCAGTTTACTAGTGCATTTGT 539
 DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180
 QY 540 TCAGTGTTCGTAGGGCTTTCCCGACTGTTGGCTTTCAAGTTATATGGATGATGTTGTA 599
 DB 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200
 QY 600 TTGGGGCGAAGTCTGTACAACTCTTGAGTCCCTTTTACCTTATACCAATTTCTT 659
 DB 200 rTrpGlyProSerLeuTyraAsnIleLeuAsnProPheLeuProLeuLeuProIlePhePh 220
 QY 660 TTGCTTTGGGTATACATT 678
 DB 220 eCysLeuTrpValTyriIle 226

RESULT 12

JQ2115

surface antigen - hepatitis B virus (subtype adr, strain p1WK146)

C;Species: hepatitis B virus, HBV

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004

C;Accession: JQ2115

R;Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus

J. Gen. Virol. 74, 1341-1348, 1993

A;Title: Genetic relatedness of hepatitis B virus strains of diverse geographical origin

A;Reference number: JQ2044; MUID:93329382; PMID:8336122

A;Contents: genogroup C

A:Accession: JQ2115
A:Molecule type: DNA
A:Residues: 1-226 <NOR>
A:Cross-references: UNIPROT:Q9QM14; UNIPARC:UPI0000178437
C:Genetics:
A:Gene: S
C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 5,21e-99 Length: 226
Score: 1192.00 Matches: 217
Percent Similarity: 96.04% Conservative: 1
Best Local Similarity: 95.59% Mismatches: 8
Query Match: 93.27% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x JQ2115 (1-226)

```
QY 1 ATGGAGAACACAAATCCTAGGATTCCTAGGACCCCTGCTCGTGTATACAGCGGGGTTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20

QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGACTC-TGGTGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 40

QY 120 TTTCTAGGGGGAGCACCCACGCTGTTCTCGGCCAAATTCGCAGTCCCAACCTCCAATCA 179
DB 41 PheLeuGlyGlyAlaProThrCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60

QY 180 CTCACCAACCTCTTGCTCCCAATTTGTCCTGGCTATCGCTGAGTGTCTCGGGGTTT 239
DB 60 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80

QY 240 TATCATATTCTTCATCTCTGCTATGCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 299
DB 80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100

QY 300 CCAAGGTATGTGGCCGTTGTCCTTACTTCCAGGAACATCAACCCAGCACCGGGGCC 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120

QY 360 ATGCAAGACTGACAGACTCTCTGCTCAAGAACTCTACGTTTCCTCTCTCTCTCTCTCT 419
DB 120 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140

QY 420 AAAACCTTCGGAGCGGAACTGATCTGTATTCCTCCATCCCATCTCTGGCTTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheAlaLy 160

QY 480 ATTCTTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT 539
DB 160 sPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuValProPheVa 180

QY 540 TCAGTGGTTCGTAGGCTTTCCCGGCTTTCCCGGCTTTTCAGTTATATGATGATGGTA 599
DB 180 lGlnTrpPheValGlyLeuSerProThrValTrpProSerAlaIleTrpMetMetTrpTy 200

QY 600 TTGGGGCGAAGTCTGTACATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
DB 200 rTrpGlyProSerLeuTyrAsnIleLeuSerProPheLeuProLeuLeuProIlePhePh 220

QY 660 TTGTCTTTGGGTATACATT 678
DB 220 eCysLeuTrpValTyrlle 226
```

RESULT 13
JQ2094

surface antigen - hepatitis B virus (subtype adr-, strain Wan)

C:Species: hepatitis B virus, HBV

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 31-Dec-2004

C:Accession: JQ2094; PQ0604

R:Order, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Muehahwar, I.K.; Magnius,

J. Gen. Virol. 74, 1341-1348, 1993

A:Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin

A:Reference number: JQ2044; MUID:93329382; PMID:8336122

A:Contents: genogroup C

A:Accession: JQ2094

A:Molecule type: DNA

A:Residues: 1-226 <NOR>

A:Cross-references: UNIPROT:Q9B6S4; UNIPROT:Q9DH50; UNIPROT:Q9DUK5; UNIPROT:Q9B8K9; UNIP

PROT:Q9PXA0; UNIPROT:Q9DKP5; UNIPROT:Q917V3; UNIPROT:Q91557; UNIPROT:Q99HU8; UNIPROT:Q9E

DH0; UNIPROT:Q9VIM3; UNIPROT:Q9DTC4; UNIPROT:Q9DKQ2; UNIPROT:Q918J5; UNIPROT:Q91560; UN

R:Order, H.; Courouce, A.M.; Magnius, L.O.

J. Gen. Virol. 73, 3141-3145, 1992

A:Title: Molecular basis of hepatitis B virus serotype variations within the four major

A:Reference number: PQ0453; MUID:93107848; PMID:1469353

A:Accession: PQ0604

A:Molecule type: DNA

A:Residues: 101-180 <NO2>

A:Cross-references: UNIPARC:UPI0000178458

C:Genetics:

A:Gene: S

C:Keywords: surface antigen

Alignment Scores:
Pred. No.: 6,41e-99 Length: 226
Score: 1191.00 Matches: 216
Percent Similarity: 95.59% Conservative: 1
Best Local Similarity: 95.15% Mismatches: 9
Query Match: 93.13% Indels: 2
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_155_835 (1-681) x JQ2094 (1-226)

```
QY 1 ATGGAGAACACAAATCCTAGGATTCCTAGGACCCCTGCTCGTGTATACAGCGGGGTTTTTC 60
DB 1 MetGluAsnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe 20

QY 61 TCGTTGACAAAGATCCTCAATACCGCAGAGTCTAGACTC-TGGTGACTTCTCTCAAT 119
DB 21 LeuLeuThrArgIleLeuThrIleProGlnSerLeuAspSerTrpTrpThrSerLeuAen 40

QY 120 TTTCTAGGGGGAGCACCCACGCTGTTCTCGGCCAAATTCGCAGTCCCAACCTCCAATCA 179
DB 41 PheLeuGlyGlyAlaProValCys-ProGlyGlnAsnSerGlnSerProThrSerAsnHi 60

QY 180 CTCACCAACCTCTTGCTCCCAATTTGTCCTGGCTATCGCTGAGTGTCTCGGGGTTT 239
DB 60 sSerProThrSerCysProProIleCysProGlyTyrArgTrpMetCysLeuArgArgPh 80

QY 240 TATCATATTCTTCATCTCTGCTATGCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 299
DB 80 eileilePheLeuPheileLeuLeuLeuCysLeuilePheLeuLeuValLeuLeuAspTy 100

QY 300 CCAAGGTATGTGGCCGTTGTCCTTACTTCCAGGAACATCAACCCAGCACCGGGGCC 359
DB 100 rGlnGlyMetLeuProValCysProLeuLeuProGlyThrSerThrThrSerThrGlyPr 120

QY 360 ATGCAAGACTGACAGACTCTCTGCTCAAGAACTCTACGTTTCCCTCTCTCTCTCTCTCT 419
DB 120 oCysIysThrCysThrIleProAlaGlnGlyThrSerMetPheProSerCysCysTh 140

QY 420 AAAACCTTCGGAGCGGAACTGATCTGTATTCCTCCATCCCATCTCTGGCTTTCGCAAG 479
DB 140 rLysProSerAspGlyAsnCysThrCysIleProIleProSerSerTrpAlaPheValAr 160

QY 480 ATTCTTATGGAGTGGGCTCAGTCCGTTTCTCTGGCTCAGTTTACTAGTGCATTTGT 539
DB 160 gPheLeuTrpGluTrpAlaSerValArgPheSerTrpLeuSerLeuLeuAlaProPheVa 180

QY 540 TCAGTGGTTCGTAGGCTTTCCCGGCTTTCCCGGCTTTTCAGTTATATGATGATGGTA 599
DB 180 lGlnTrpPheValGlyLeuSerProThrValTrpLeuSerValIleTrpMetMetTrpTy 200

QY 600 TTGGGGCGGAGACTGTGTACAAATCTTGAGTCCCTTTTACCTCTATTACCAATTTCTT 659
```

420 A A A A C C T T C G G A C G G A A A C T G C A C T T G T A T T C C C A T C C C A T C A T C T C G G G C T T T C G C A A G 479

140 r l y s p r o s e r a s p g l y a s n c y s t h r c y l i e p r o l l e p r o s e r s e r t r p a l a p h e a l a a r 160

480 A T T C C T A T G G G A G T G G C C T C A G T C C G T T T C T C T G C C T C A G T T T A C T A G T A G C C A T T T G T 539

160 g p h e l e u t r p g l u t r p a l a s e r a l a a r g p h e s e r t r p l e u s e r l e u l e u v a l p r o p h e v a 180

540 T C A G T G T T C G T A G G G C T T T C C C C C A C T G T T T G G C T T T C A G T A T A T A T G G A T G A T G T G T A 599

180 l g l n t r p p h e v a l g l y l e u s e r p r o t h r v a l t r p l e u s e r v a l l e t r p m e t t r p t y 200

600 T T C G G G C G C A A G T C T G T A C A C A C A T T T G A G T C C C T T T T T A C T C T A T T A C C A A T T T T C T T 659

200 r t r p g l y p r o s e r l e u h i s a n l e u a s n p r o p h e l e u p r o l l e p h e p h 220

660 T T C T C T T T G G G T A T A C A T T 678

220 e c y s l e u t r p v a l t y r i l e 226

RESULT 15

T13469

N;large surface antigen - hepatitis B virus (isolate 09D09HCC)

N;Alternate names: envelope protein

N;Contains: major surface antigen; middle surface antigen

C;Species: hepatitis B virus, HBV

A;Variety: isolate 09D09HCC

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C;Accession: T13469

R;Takahashi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.

Arch. Virol. 143, 2313-2326, 1998

A;Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients

A;Reference number: Z17684; MUID:99129050; PMID:9930189

A;Accession: T13469

A;Molecule type: DNA
A;Residues: 1-384 <YAK>
A;Cross-references: UNIPROT:Q9YZU2; UNIPARC:UPI00000FIC85; EMBL:AB014368; NID:
A;Experimental source: Japanese patient with hepatocellular carcinoma isolate
C;Genetics:
A;Gene: S
A;Introns: 123/2
A;Superfamily: hepatitis B virus surface antigen
C;Keywords: glycoprotein; surface antigen

Alignment Scores:	7.75e-99	Length:	384
Pred. No.:	1190.00	Matches:	217
Score:	1190.00	Conservative:	9
Percent Similarity:	95.59%	Mismatches:	0
Best Local Similarity:	95.59%	Indels:	2
Query Match:	93.11%	Gaps:	0
DB:	2		

US-10-761-006A-1_COPY_155_835 (1-681) x T13469 (1-384)

Qy	1	ATGAGAACACAAATCCTCAGAGATTCCTAGGACCCCTGCTCGTGTGTACAGCGGGGTTTTTC	60
Db	159	MetGluAnThrThrSerGlyPheLeuGlyProLeuLeuValLeuGlnAlaGlyPhePhe	178
Qy	61	TCGTTGACAGAATCCTCACAAATCCGACAGTCTAGACTC-TGGTGGACTTCTCTCAAT	119
Db	179	LeuLeuThrArgIleLeuThrIleProGlnSerLeuAepSerTrpTrpThrSerLeuAen	198
Qy	120	TTTCTAGGGGAGACCCACGTTGTTCTGGCCAAATTCGCAGTCCCCACCTCCAATCA	179
Db	199	PheLeuGlyGlyAlaPro**Cys-ProGlyGlnAenSerGlnSerProThrSerAenHi	218
Qy	180	CTACCAACCTTTGCTCTCCAAATTGTCCTGGCTATCGCTGGATGTGTCGCGCGTTT	239
Db	218	sSerProThrSerCysProProIleCysProGlyYrZargIrpMetCysLeuAerGargh	238
Qy	240	TATCATATTCCTCTTCATCTCGCTGCATGCTCATCTTCTTGTGGTTCTTCTCGGACTA	299

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:40:32 ; Search time 18.262 Seconds
(without alignments)
3320.243 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595
Perfect score: 139
Sequence: 1 ACACCTCTGCTCAAGAAA.....AACCTTCGACGGAAACTGC 69

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10761006/runat_27122005_192900_15743/app.query.fasta_1.1102
-DB-A=Geneseq -QFWT=fastan -SUPPLX=rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006@cgn 1 1 476 @runat_27122005_192900_15743 -NCPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	100.0	400	3 AAY54045	AAY54045 Amino aci
2	133	95.7	38	5 AAE17024	AAE17024 Hepatitis
3	133	95.7	48	5 AAE17025	AAE17025 Hepatitis
4	130	93.5	24	2 AAR50965	AAR50965 Peptide c
5	130	93.5	24	2 AAR50961	AAR50961 Amino aci
6	130	93.5	24	2 AAR50968	AAR50968 Peptide c
7	130	93.5	24	2 AAR94433	AAR94433 HBV surfa
8	130	93.5	24	2 AAY04027	AAY04027 HBsAg pep
9	130	93.5	24	2 AAE25520	AAE25520 Hepatitis

10	130	93.5	43	2	AAR11409	Hepatitis
11	130	93.5	51	9	ADW68797	Hepatitis
12	130	93.5	55	9	ADW68794	Hepatitis
13	130	93.5	60	9	ADW68793	Hepatitis
14	130	93.5	70	9	ADW68806	Hepatitis
15	130	93.5	70	9	ADW68821	Hepatitis
16	130	93.5	70	9	ADW68856	Hepatitis
17	130	93.5	71	2	AAW65505	Hepatitis
18	130	93.5	71	3	AAW85118	Hepatitis
19	130	93.5	75	9	ADW68792	Hepatitis
20	130	93.5	90	9	ADW68815	Hepatitis
21	130	93.5	90	9	ADW68814	Hepatitis
22	130	93.5	154	9	ADW68791	Hepatitis
23	130	93.5	154	9	ADW68813	Hepatitis
24	130	93.5	160	7	ADG76930	Hepatitis
25	130	93.5	165	7	ADG76927	Hepatitis
26	130	93.5	209	2	AAR55288	Deduced s
27	130	93.5	226	1	AAP30096	Sequence
28	130	93.5	226	1	AAP30060	Surface a
29	130	93.5	226	1	AAP60223	Sequenc o
30	130	93.5	226	1	AAP80417	Sequence
31	130	93.5	226	2	AAR33252	HBsAg enc
32	130	93.5	226	2	AAR55282	Deduced s
33	130	93.5	226	3	AAB35766	Peptidyl
34	130	93.5	226	4	AAE04711	Hepatitis
35	130	93.5	226	4	AAB98041	Hepatitis
36	130	93.5	226	4	AAB80956	Viral pro
37	130	93.5	226	5	AAM52694	Hepatitis
38	130	93.5	226	5	ABB05219	Hepatitis
39	130	93.5	226	5	AAU97646	Hepatitis
40	130	93.5	226	8	ADU69200	Hepatitis
41	130	93.5	226	9	ADW68808	Hepatitis
42	130	93.5	226	9	ADY62464	Hepatitis
43	130	93.5	226	9	AEA08824	Hepatitis
44	130	93.5	250	2	AAR11496	Hepatitis
45	130	93.5	251	2	AAR11495	RPI35/HBS

ALIGNMENTS

RESULT 1

AAV54045
ID AAY54045 standard; protein; 400 AA.

XX AAY54045;

XX 27-MAR-2000 (first entry)

XX Amino acid sequence of a HBV large surface antigen protein.

XX HBV; HBV surface antigen-'S'-133 Oon strain (Met to Thr); DNA polymerase;
XX large surface antigen; core protein; transactivating X protein;
XX hepatitis vaccine; HBV infection; hepatocellular carcinoma.

XX Hepatitis B virus.

XX Key Location/Qualifiers

FT Misc-difference 14 /note= "encoded by AGC"

FT Misc-difference 174..400 /note= "these residues are specifically claimed in claim

FT 29"

FT Misc-difference 298..320 /note= "these residues are specifically claimed in claim

FT 23"

XX WO9966048-A1.

XX 23-DEC-1999.

XX 19-JUN-1998; 98WO-SG000046.

XX 19-JUN-1998; 98WO-SG000046.

XX (COVE-) GOVERNMENT REPUBLIC SINGAPORE.
 PA Oon CJ, Lim GK, Zhao Y, Chen WN;
 PI WPI; 2000-106104/09.
 XX DR N-PSDB; AA237088.
 XX PT New isolated hepatitis B virus strain, useful for, e.g. treatment of
 PT hepatitis infection.
 XX PS Claim 23; Page 39-40; 68pp; English.
 XX The present sequence is encoded by the genome of an isolated strain of
 CC Hepatitis B virus designated human Hepatitis B virus (HBV) surface
 CC antigen-'S'-133 On strain (Met to Thr). The viral genome is deposited as
 CC ECCC accession numbers P97121501, P97121502 and P97121503. The nucleotide
 CC sequence was isolated from hepatocellular carcinoma (HCC). The nucleotide
 CC sequence encodes four overlapping proteins, which are a DNA polymerase, a
 CC large surface antigen, a core protein, and a transactivating X protein.
 CC The large surface antigen differs from the wild type sequence in that it
 CC contains a Thr at position 133 of the wild type sequence instead of a
 CC Met. The proteins are used to produce antibodies. The proteins,
 CC polynucleotide and antibodies can be used for detecting the novel HBV
 CC strain. The HBV polypeptides can also be used in hepatitis vaccines. The
 CC HBV novel strain polypeptides can be used to identify compounds for
 CC treating or preventing HBV infection or hepatocellular carcinoma
 XX Sequence 400 AA;
 SQ

Alignment Scores:
 Pred. No.: 6,24e-12 Length: 400
 Score: 139.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAY54045 (1-400)
 QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTGTGTCTGTACAAACCTTCGGAC 60
 Db 299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318
 QY 61 GGAAGACTGC 69
 Db 319 GlyAsnCys 321

RESULT 2
 AAEL17024
 ID AAEL17024 standard; peptide; 38 AA.
 XX AC AAEL17024;
 XX DT 18-APR-2002 (first entry)
 XX Hepatitis B virus (HBV) adw subtype S antigen #1.
 XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
 KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
 KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
 KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
 KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
 KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
 KW antiprotozoal; S antigen.
 XX Hepatitis B virus.
 OS WO200198333-A2.
 XX 27-DEC-2001.
 XX 22-JUN-2001; 2001WO-GB002817.

XX 22-JUN-2000; 2000GB-00015308.
 PR 06-OCT-2000; 2000GB-00024544.
 XX PA (CELL-) CELLTech PHARM LTD.
 XX PI Page M, Li J, Pumpens P;
 XX WPI; 2002-098223/13.
 XX DR N-PSDB; AA27425.
 XX PT New proteins comprising a modified hepatitis B core antigen, useful as a
 PT vaccine in prophylactic or therapeutic vaccination of the human or animal
 PT body, particularly against hepatitis B virus infection.
 XX Disclosure; Fig 3B; 40pp; English.
 XX The invention relates to modified proteins comprising hepatitis B virus
 CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
 CC repeats has been deleted and the protein comprising the C-terminal
 CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
 CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
 CC present the epitope to the immune system. This chimeric protein or its
 CC nucleic acid is useful as a vaccine or in a method of prophylactic or
 CC therapeutic vaccination of the human or animal body, particularly against
 CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
 CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
 CC also be used as the basis of a prophylactic vaccine against a range of
 CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
 CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
 CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
 CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
 CC diarrhoea, meningitis or gonorrhea. The present sequence is Hepatitis B
 CC virus (HBV) adw subtype S antigen
 XX Sequence 38 AA;
 SQ

Alignment Scores:
 Pred. No.: 4,27e-11 Length: 38
 Score: 133.00 Matches: 22
 Percent Similarity: 95.65% Conservative: 0
 Best Local Similarity: 95.65% Mismatches: 1
 Query Match: 95.68% Indels: 0
 DB: 5 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAEL17024 (1-38)
 QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTGTGTCTGTACAAACCTTCGGAC 60
 Db 16 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProSerAsp 35
 QY 61 GGAAGACTGC 69
 Db 36 GlyAsnCys 38

RESULT 3
 AAEL17025
 ID AAEL17025 standard; peptide; 48 AA.
 XX AC AAEL17025;
 XX DT 18-APR-2002 (first entry)
 XX Hepatitis B virus (HBV) adw subtype S antigen #2.
 XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
 KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
 KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
 KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
 KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
 KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
 KW antiprotozoal; S antigen.

OS Hepatitis B virus.
XX WO200198333-A2.
XX
XX
XX 27-DEC-2001.
XX
XX
XX 22-JUN-2001; 2001WO-GB002817.
XX
XX 22-JUN-2000; 2000GB-00015308.
XX 06-OCT-2000; 2000GB-00024544.
XX
XX (CELL-) CELLTech PHARM LTD.
XX
XX
XX Page M, Li J, Pumpens P;
XX
XX WPI; 2002-098223/13.
XX DR N-PSDB; AAD27426.
XX
XX New proteins comprising a modified hepatitis B core antigen, useful as a
PT vaccine in prophylactic or therapeutic vaccination of the human or animal
PT body, particularly against hepatitis B virus infection.
XX
XX Disclosure; Fig 3C; 40pp; English.
XX
XX The invention relates to modified proteins comprising hepatitis B virus
CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC repeats has been deleted and the protein comprising the C-terminal
CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC present the epitope to the immune system. This chimeric protein or its
CC nucleic acid is useful as a vaccine or in a method of prophylactic or
CC therapeutic vaccination of the human or animal body, particularly against
CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC also be used as the basis of a prophylactic vaccine against a range of
CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC diarrhoea, meningitis or gonorrhoea. The present sequence is Hepatitis B
CC virus (HBV) adw subtype S antigen
XX
SQ Sequence 48 AA;

Alignment Scores:
Pred. No.: 4.33e-11 Length: 48
Score: 133.00 Matches: 22
Percent Similarity: 95.65% Conservative: 0
Best Local Similarity: 95.65% Mismatches: 1
Query Match: 95.68% Indels: 0
DB: 5 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAE17025 (1-48)
QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 16 ThrThrProAlaGlnGlyAenSerMetPheProSerCysCysThrLysProSerAsp 35

QY 61 GGAACCTGC 69
Db 36 GlyAenCys 38

RESULT 4
AAR50965
ID AAR50965 standard; peptide; 24 AA.
XX
XX AAR50965;
XX
XX 25-MAR-2003 (revised)
DT 14-OCT-1994 (first entry)
XX
XX Peptide cross reactive with anti-HBsAg antiserum.
DE
XX

KW Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
KW vaccine; treatment; antiserum; adw; surface antigen; viral infection.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 10
XX /label= Other.
XX /note= "Met (O)."
XX
XX WO9405698-A1.
XX
XX 17-MAR-1994.
XX
XX 30-AUG-1993; 93WO-EP002342.
XX
XX 01-SEP-1992; 92AT-00001746.
XX
XX (UNNA-) UNITED NATIONS IND DEV ORG.
XX Manivel V, Rao KV, Panda SK;
XX WPI; 1994-101123/12.
XX
XX Peptide(s) for use in diagnosis or vaccines for hepatitis B virus -
PT comprising oligomers of an epitope sequence of hepatitis B surface
PT antigen.
XX
XX Claim 2; Page 14; 26pp; English.
XX
XX The peptide shows cross reactivity with anti-hepatitis B surface antigen
CC (HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a
CC group specific determinant of HBsAg. It can be used for the diagnosis of
CC hepatitis B virus infection and in vaccines against the virus,
CC particularly in combination with alum as adjuvant and with a myristic
CC acid residue added to the amino terminus of the peptide. This sequence is
CC an analogue of the sequence described in AAR50961. See also in AAR50962-
CC 70. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 24 AA;

Alignment Scores:
Pred. No.: 1.17e-10 Length: 24
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAR50965 (1-24)
QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 2 ThrThrProAlaGlnGlyAenSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAACCTGC 69
Db 22 GlyAenCys 24

RESULT 5
AAR50961
ID AAR50961 standard; peptide; 24 AA.
XX
XX AAR50961;
XX
XX 25-MAR-2003 (revised)
DT 14-OCT-1994 (first entry)
XX
XX Amino acids 124-147 of Hepatitis B surface antigen subtype adw.
DE
XX Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
KW vaccine; treatment; antiserum; adw; surface antigen; viral infection.
KW
XX

OS	Synthetic.
XX	
PN	WO9405698-A1.
XX	
PD	17-MAR-1994.
XX	
PF	30-AUG-1993; 93WO-EP002342.
XX	
PR	01-SEP-1992; 92AT-00001746.
XX	
PA	(UNNA-) UNITED NATIONS IND DEV ORG.
XX	
PI	Manivel V, Rao KV, Panda SK;
XX	
DR	WPI; 1994-101123/12.
XX	
PT	Peptide(s) for use in diagnosis or vaccines for hepatitis B virus -
PT	comprising oligomers of an epitope sequence of hepatitis B surface
PT	antigen.
XX	
PS	Claim 2; Page 14; 26pp; English.
XX	
CC	The peptide shows cross reactivity with anti-hepatitis B surface antigen
CC	(HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a
CC	group specific determinant of HBsAg. It can be used for the diagnosis of
CC	hepatitis B virus infection and in vaccines against the virus,
CC	particularly in combination with alum as adjuvant and with a myristic
CC	acid residue added to the amino terminus of the peptide. Analogues of
CC	this sequence are described in AAR50962-70. (Updated on 25-MAR-2003 to
CC	correct PN field.)
XX	
SQ	Sequence 24 AA;
Alignment Scores:	
Pred. No.:	1.17e-10 Length: 24
Score:	130.00 Matches: 21
Percent Similarity:	95.65% Conservative: 1
Best Local Similarity:	91.30% Mismatches: 1
Query Match:	93.53% Indels: 0
DB:	2 Gaps: 0
US-10-761-006A-1_COPY_527_595 (1-69) x AAR50961 (1-24)	
QY	1 ACAGACTCTGCTCAAGAACTCTACGTTCCTCTTGTTGGTGTACAAAACCTTCGGAC 60
Db	2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProthrAsp 21
QY	61 GGAAACTGC 69
Db	22 GlyAsnCys 24
RESULT 6	
AAR50968	
ID	AAR50968 standard; peptide; 24 AA.
XX	
AC	AAR50968;
XX	
DT	25-MAR-2003 (revised)
DT	14-OCT-1994 (first entry)
XX	
DE	Peptide cross reactive with anti-HBsAg antiserum.
XX	
KW	Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
KW	vaccine; treatment; antiserum; adv; surface antigen; viral infection.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 18
FT	/label= MeLys.
XX	
PN	WO9405698-A1.
XX	

PD	17-MAR-1994.
XX	
PF	30-AUG-1993; 93WO-EP002342.
XX	
PR	01-SEP-1992; 92AT-00001746.
XX	
PA	(UNNA-) UNITED NATIONS IND DEV ORG.
XX	
PI	Manivel V, Rao KV, Panda SK;
XX	
DR	WPI; 1994-101123/12.
XX	
PT	Peptide(s) for use in diagnosis or vaccines for hepatitis B virus -
PT	comprising oligomers of an epitope sequence of hepatitis B surface
PT	antigen.
XX	
PS	Claim 2; Page 14; 26pp; English.
XX	
CC	The peptide shows cross reactivity with anti-hepatitis B surface antigen
CC	(HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a
CC	group specific determinant of HBsAg. It can be used for the diagnosis of
CC	hepatitis B virus infection and in vaccines against the virus,
CC	particularly in combination with alum as adjuvant and with a myristic
CC	acid residue added to the amino terminus of the peptide. Analogues of
CC	this sequence are described in AAR50962-70. (Updated on 25-MAR-2003 to
CC	correct PN field.)
XX	
SQ	Sequence 24 AA;
Alignment Scores:	
Pred. No.:	1.17e-10 Length: 24
Score:	130.00 Matches: 21
Percent Similarity:	95.65% Conservative: 1
Best Local Similarity:	91.30% Mismatches: 1
Query Match:	93.53% Indels: 0
DB:	2 Gaps: 0
US-10-761-006A-1_COPY_527_595 (1-69) x AAR50968 (1-24)	
QY	1 ACAGACTCTGCTCAAGAACTCTACGTTCCTCTTGTTGGTGTACAAAACCTTCGGAC 60
Db	2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProthrAsp 21
QY	61 GGAAACTGC 69
Db	22 GlyAsnCys 24
RESULT 7	
AAR94433	
ID	AAR94433 standard; peptide; 24 AA.
XX	
AC	AAR94433;
XX	
DT	27-AUG-2003 (revised)
DT	31-MAY-1996 (first entry)
XX	
DE	HBV surface protein B-cell epitope.
XX	
KW	Molecular presentation; FHV; virion-like particle; capsid protein;
KW	capsomer; RNA-2 gene; epitope; HBV; vaccine.
XX	
OS	Hepatitis B virus.
XX	
PN	WO9605293-A1.
XX	
PD	22-FEB-1996.
XX	
PF	04-AUG-1995; 95WO-EP003114.
XX	
PR	08-AUG-1994; 94AT-00001545.
XX	
PA	(UNNA-) UNITED NATIONS IND DEV ORG.
XX	

Synthetic.
WO9405698-A1.
17-MAR-1994.
30-AUG-1993; 93WO-EP002342.
01-SEP-1992; 92AT-00001746.
(UnNA-) UNITED NATIONS IND DEV ORG.
Manivel V, Rao KV, Panda SK;
WPI; 1994-101123/12.
Peptide(s) for use in diagnosis or vaccines for hepatitis B virus - comprising oligomers of an epitope sequence of hepatitis B surface antigen.
Claim 2; Page 14; 26pp; English.
The peptide shows cross reactivity with anti-hepatitis B surface antigen (HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a group specific determinant of HBsAg. It can be used for the diagnosis of hepatitis B virus infection and in vaccines against the virus, particularly in combination with alum as adjuvant and with a myristic acid residue added to the amino terminus of the peptide. Analogues of this sequence are described in AAR50962-70. (Updated on 25-MAR-2003 to correct PN field.)

Alignment Scores:
Pred. No.: 1.17e-10 Length: 24
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAR50961 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTGTGGTGTACAAAACCTTCGGAC 60
|||||
Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProthrAsp 21
|||||

QY 61 GGAAACTGC 69
|||||
Db 22 GlyAsnCys 24
|||||

RESULT 6
AAR50968
ID AAR50968 standard; peptide; 24 AA.
XX AC AAR50968;
XX DT 25-MAR-2003 (revised)
DT 14-OCT-1994 (first entry)
XX DE Peptide cross reactive with anti-HBsAg antiserum.
XX KW Hepatitis B virus; HBV; epitope; antigen; HBsAg; diagnosis; detection;
KW vaccine; treatment; antiserum; adv; surface antigen; viral infection.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 18
FT /label= MeLys.
XX PN WO9405698-A1.
XX

PD 17-MAR-1994.
XX 30-AUG-1993; 93WO-EP002342.
XX 01-SEP-1992; 92AT-00001746.
XX (UNNA-) UNITED NATIONS IND DEV ORG.
XX Manivel V, Rao KV, Panda SK;
XX WPI; 1994-101123/12.
XX Peptide(s) for use in diagnosis or vaccines for hepatitis B virus - comprising oligomers of an epitope sequence of hepatitis B surface antigen.
XX Claim 2; Page 14; 26pp; English.
XX The peptide shows cross reactivity with anti-hepatitis B surface antigen (HBsAg) antiserum. The peptide comprises the "a" epitope of HBsAg, a group specific determinant of HBsAg. It can be used for the diagnosis of hepatitis B virus infection and in vaccines against the virus, particularly in combination with alum as adjuvant and with a myristic acid residue added to the amino terminus of the peptide. This sequence is an analogue of the sequence described in AAR50961. See also in AAR50962-70. (Updated on 25-MAR-2003 to correct PN field.)

Alignment Scores:
Pred. No.: 1.17e-10 Length: 24
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAR50968 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTGTGGTGTACAAAACCTTCGGAC 60
|||||
Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProthrAsp 21
|||||

QY 61 GGAAACTGC 69
|||||
Db 22 GlyAsnCys 24
|||||

RESULT 7
AAR94433
ID AAR94433 standard; peptide; 24 AA.
XX AC AAR94433;
XX DT 27-AUG-2003 (revised)
DT 31-MAY-1996 (first entry)
XX DE HBV surface protein B-cell epitope.
XX KW Molecular presentation; FHV; virion-like particle; capsid protein;
KW capsomer; RNA-2 gene; epitope; HBV; vaccine.
XX OS Hepatitis B virus.
XX XX WO9605293-A1.
XX PD 22-FEB-1996.
XX PF 04-AUG-1995; 95WO-EP003114.
XX PR 08-AUG-1994; 94AT-00001545.
XX PA (UNNA-) UNITED NATIONS IND DEV ORG.
XX

PI Baralle PE, Scodeller E, Tisminetzky S;
 XX WPI; 1996-139691/14.
 XX
 XX New molecular presentation system - comprising a viral protein from a
 PT small insect virus in which heterologous amino acid sequences are
 PT inserted.
 XX
 XX Disclosure; Page 7; 39pp; English.
 XX
 XX The B-cell epitope (AAR94433) of the hepatitis B virus surface protein
 CC may be inserted into the outward-directed L1, L2, L3, L2 or L3 loop of
 CC the Flock House virus capsid protein (AAR8755). The capsid protein
 CC provides a conformationally suitable location for this (or other, see
 CC AAR94430-32 and AAR94434-45) heterologous peptides. A molecular
 CC presentation system is obtd. The FHV recombinant capsomer can be
 CC expressed in E. coli. Alternatively, expression in insect cells using a
 CC baculovirus vector results in prodn. of mature virus-like particles.
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 XX Sequence 24 AA;

Alignment Scores:
 Pred. No.: 1.17e-10 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x AAR94433 (1-24)

QY 1 ACGACTCTCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTACAAACCTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProThrAsp 21

QY 61 GGAAACTGC 69
 Db 22 GlyAsnCys 24

RESULT 8
 AAY04027
 ID AAY04027 standard; peptide; 24 AA.
 AC AAY04027;

XX
 XX 17-OCT-2003 (revised)
 DT 20-DEC-1999 (first entry)
 XX
 XX HBsAg peptide 124-147 for use in immunogenic complex.
 DE
 XX Macroglobulin; immunogenic complex; hepatitis; surface antigen; HBsAg;
 KW antigen; vaccine.
 XX

OS Hepatitis B virus.

XX W09950303-A2

PN 07-OCT-1999.

XX 01-APR-1999; 99WO-US007236.

XX 01-APR-1998; 98US-00053301.

XX 31-MAR-1999; 99US-00282826.

XX (UYDU-) UNIV DUKE.

XX Pizzo SV, Gron H;

XX WPI; 1999-601324/51.

XX Immune response modulator alpha-2 macroglobulin complex useful in the
 PT treatment of HIV or hepatitis.

XX Claim 3; Page 77; 103pp; English.

XX A stable complex comprising at least one intact biomolecule and activated
 CC alpha-2-macroglobulin (alpha-2M) having an intact bait region is claimed,
 CC in which the biomolecule is covalently bound to an amino acid residue of
 CC a cleaved thiol ester of the alpha-2M, the amino acid residue being a
 CC glutamyl residue and/or a cysteinyl residue. Preparation of the complex
 CC comprises activation of alpha-2M with a nucleophilic compound followed by
 CC incubation with the intact biomolecule at elevated temperature. Covalent
 CC incorporation of the intact biomolecule is thus effected without use of a
 CC protease. The obtained complex may be used as an antigen for stimulating
 CC immune response, e.g. in the form of a vaccine. Enhanced presentation of
 CC the biomolecule is provided, especially for those antigens which are
 CC poorly immunogenic. Reduction of immunodominance of particular epitopes
 CC is also provided. Preferably the biomolecule is an HIV antigen or a
 CC hepatitis virus antigen, or a peptide, fragment, hybrid or chimeric
 CC peptide thereof. The present sequence is a preferred example of a
 CC hepatitis peptide which may be used in the complex. It represents
 CC residues 124-147 of Hepatitis B virus surface antigen (HBsAg) (Updated on
 CC 17-OCT-2003 to standardise OS field)

XX Sequence 24 AA;

Alignment Scores:
 Pred. No.: 1.17e-10 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x AAY04027 (1-24)

QY 1 ACGACTCTCTCAAGGAACTCTACGTTTCCCTCTTGTGTGTACAAACCTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrllysProThrAsp 21

QY 61 GGAAACTGC 69
 Db 22 GlyAsnCys 24

RESULT 9
 AAE25520
 ID AAE25520 standard; peptide; 24 AA.
 AC AAE25520;

XX
 XX 06-AUG-2003 (revised)
 DT 04-NOV-2002 (first entry)

XX Hepatitis b virus surface antigen (HBsAg) peptide.

XX Stable complex; alpha-2-macroglobulin; immunocompetence; vaccine;
 KW immune response; macrophage; tumour; human immunodeficiency virus; HIV;
 KW cytostatic; antimicrobial; alpha2M; anti-HIV; virucide; HBsAg.

OS Hepatitis B virus.

XX US6403032-B1

XX 11-JUN-2002.

XX 31-MAR-1999; 99US-00282826.

XX 01-APR-1998; 98US-00053301.

XX (UYDU-) UNIV DUKE.

XX Pizzo S, Gron H;

XX WPI; 1999-601324/51.

PT Immune response modulator alpha-2 macroglobulin complex useful in the
PT treatment of HIV or hepatitis.
XX
PS Claim 3; Col 35; 34pp; English.
XX
CC The invention relates to a stable complex comprising an intact
CC biomolecule with a nucleophilic group which is covalently bound to
CC activated alpha-2-macroglobulin (alpha2M) with an intact bait region. The
CC stable complex is useful for enhancing host immunocompetence and for
CC preparing vaccines for preventing and treating disease states. It is
CC useful for increasing immune response to a poorly immunogenic antigen,
CC and for suppressing immune response to a particular antigen. It is most
CC preferably useful for administering antigens recognised by the macrophage
CC in view of the existence on the macrophage of receptors for alpha2M. It
CC is useful for treating cancer or infection, where the isolated protein is
CC derived from tumour, abnormal cells or infectious organism, and this
CC protein can be used as an antigen and prepared in alpha2M complex which
CC is then administered in the form of vaccine to immunise mammalian
CC patients in need of such treatment. It can also be used for inducing
CC immune response against human immunodeficiency virus (HIV) antigens, and
CC hepatitis virus antigens. The present sequence is Hepatitis B virus
CC surface antigen (HBsAg) peptide used in the exemplification of the
CC invention. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 24 AA;

Alignment Scores:
Pred. No.: 1.17e-10 Length: 24
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAE25520 (1-24)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAAACTGC 69
Db 22 GlyAsnCys 24

RESULT 10
AAR11409
ID AAR11409 standard; protein; 43 AA.
XX
AC AAR11409;
XX
DT 25-MAR-2003 (revised)
DT 10-JUN-1991 (first entry)
XX
DE Hepatitis B surface antigen pPD1-110-148 complex epitope.
XX
KW HBsAg; chimaera; hepadnavirus; vaccine; ppV-Nhe.
XX
OS Synthetic.
XX
PN 10-APR-1991.
XX
PD 10-APR-1991.
XX
PF 19-SEP-1990; 90BP-00310264.
XX
PR 19-SEP-1989; 89GB-00021172.
PR 13-APR-1990; 90GB-00017728.
XX
XX (WELL) WELLCOME FOUND LTD.
PA (BROW/) BROWN A L.
XX
PI Brown AL, Clarke BE, Rowlands DJ;
XX

DR WPI; 1991-103861/15.
DR N-PSDB; AAQ11233.
XX
PT Particles comprising chimeric hepadnavirus core antigen protein - contg.
PT foreign aminoacid sequence comprising epitope of e.g. hepatitis A or B
PT and human immuno-deficiency virus, used as vaccine.
XX
PS Example 3; Page 10; 24pp; English.
XX
CC The insert encoding this epitope was ligated into the NheI restriction
CC site of plasmid pPV-Nhe to recreate the NheI sites which code for Ala-Ser
CC residues flanking the epitope. The recombinant plasmid was transformed
CC into E.coli strain XL-1 Blue. Positive clones were cultured to high
CC density in nutrient broth and induced to express chimaeric protein
CC comprising the desired epitope by addition of IPTG. (Updated on 25-MAR-
CC 2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 43 AA;

Alignment Scores:
Pred. No.: 1.22e-10 Length: 43
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x AAR11409 (1-43)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTTCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 18 ThrThrProAlaGlnGlyAsnSerLysPheProSerCysCysThrLysProThrAsp 37

QY 61 GGAAACTGC 69
Db 38 GlyAsnCys 40

RESULT 11
ADW68797
ID ADW68797 standard; protein; 51 AA.
XX
AC ADW68797;
XX
DT 24-MAR-2005 (first entry)
XX
DE Hepatitis B virus HBsAg surface antigen fragment #7.
XX
KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
KW antiinflammatory; DNA detection; vaccine; diagnostic;
KW hepatitis B virus infection; selectable marker.
XX
OS Hepatitis B virus.
XX
PN WO2004113369-A1.
XX
PD 29-DEC-2004.
XX
PF 17-JUN-2004; 2004WO-EP006515.
XX
PR 20-JUN-2003; 2003DE-01028080.
XX
PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI Krupka U;
XX
DR WPI; 2005-066451/07.
DR N-PSDB; ADW68786.
XX
PT New oligo- or poly-peptide variants of hepatitis B surface antigen,
PT useful for diagnosis of, and vaccination against, hepatitis infection,
PT derived from strain HDB05, also related nucleic acid and antibodies.
XX
XX Claim 3; SEQ ID NO 18; 61pp; German.

XX This sequence represents novel oligo- and poly-peptides that are variants
 CC of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
 CC from a patient with inflammation of the liver but an atypical reaction
 CC pattern of hepatitis markers. PCR amplification and sequencing indicated
 CC the 5 specified mutations in the S gene, relative to adw wild-type. These
 CC mutations all caused amino acid variations: 4 in the a-determinant region
 CC and the other close to it at residue 181. The products of the invention
 CC have virucide, hepatotropic and antiinflammatory activity. The products
 CC of the invention are used to detect specific antibodies, as probes or
 CC primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
 CC and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
 CC including use in vaccination, diagnosis and screening of donated blood.
 CC The antibodies do not bind, or bind much less strongly, to the HBsAg from
 CC virus of genotype A, subtype adw. This sequence represents a fragment of
 CC Hepatitis B virus surface protein (HBsAg) corresponding to residues 115-
 XX 165.
 XX
 SQ Sequence 51 AA;

Alignment Scores:
 Pred. No.: 1.23e-10 Length: 51
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADW68797 (1-51)

QY 1 AGACCTCTGCTCAGGAACCTCTACGTTTCCTCTGTGCTGTACAAACCTCGGAC 60
 Db 11 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 30
 QY 61 GGAAACTGC 69
 Db 31 GlyAsnCys 33

RESULT 12
 ADW68794
 ID ADW68794 standard; protein; 55 AA.
 XX
 AC ADW68794;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Hepatitis B virus HBsAg surface antigen fragment #4.
 XX
 KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
 KW antiinflammatory; DNA detection; vaccine; diagnostic;
 KW hepatitis B virus infection; selectable marker.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004113369-A1.
 XX
 PD 29-DEC-2004.
 XX
 PF 17-JUN-2004; 2004WO-EP006515.
 XX
 PR 20-JUN-2003; 2003DE-01028080.
 XX
 PA (DADE-) DADE BEHRING MARBURG GMBH.
 XX
 PI Krupka U;
 XX
 DR WPI; 2005-066451/07.
 DR N-PSDB; ADW68783.
 XX

New oligo- or poly-peptide variants of hepatitis B surface antigen,
 PT useful for diagnosis of, and vaccination against, hepatitis infection,
 PT derived from strain HBD05, also related nucleic acid and antibodies.
 PT

PS Claim 3; SEQ ID NO 15; 61pp; German.
 XX
 CC This sequence represents novel oligo- and poly-peptides that are variants
 CC of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
 CC from a patient with inflammation of the liver but an atypical reaction
 CC pattern of hepatitis markers. PCR amplification and sequencing indicated
 CC the 5 specified mutations in the S gene, relative to adw wild-type. These
 CC mutations all caused amino acid variations: 4 in the a-determinant region
 CC and the other close to it at residue 181. The products of the invention
 CC have virucide, hepatotropic and antiinflammatory activity. The products
 CC of the invention are used to detect specific antibodies, as probes or
 CC primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
 CC and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
 CC including use in vaccination, diagnosis and screening of donated blood.
 CC The antibodies do not bind, or bind much less strongly, to the HBsAg from
 CC virus of genotype A, subtype adw. This sequence represents a fragment of
 CC Hepatitis B virus surface protein (HBsAg) corresponding to residues 111-
 XX 165.
 XX
 SQ Sequence 55 AA;

Alignment Scores:
 Pred. No.: 1.24e-10 Length: 55
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservative: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADW68794 (1-55)

QY 1 AGACCTCTGCTCAGGAACCTCTACGTTTCCTCTGTGCTGTACAAACCTCGGAC 60
 Db 15 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 34
 QY 61 GGAAACTGC 69
 Db 35 GlyAsnCys 37

RESULT 13
 ADW68793
 ID ADW68793 standard; protein; 60 AA.
 XX
 AC ADW68793;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Hepatitis B virus HBsAg surface antigen fragment #3.
 XX
 KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
 KW antiinflammatory; DNA detection; vaccine; diagnostic;
 KW hepatitis B virus infection; selectable marker.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004113369-A1.
 XX
 PD 29-DEC-2004.
 XX
 PF 17-JUN-2004; 2004WO-EP006515.
 XX
 PR 20-JUN-2003; 2003DE-01028080.
 XX
 PA (DADE-) DADE BEHRING MARBURG GMBH.
 XX
 PI Krupka U;
 XX
 DR WPI; 2005-066451/07.
 DR N-PSDB; ADW68782.
 XX

New oligo- or poly-peptide variants of hepatitis B surface antigen,
 PT useful for diagnosis of, and vaccination against, hepatitis infection,
 PT derived from strain HBD05, also related nucleic acid and antibodies.
 PT

```
XX PS Claim 3; SEQ ID NO 14; 61pp; German.
XX
XX This sequence represents novel oligo- and poly-peptides that are variants
XX of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
XX from a patient with inflammation of the liver but an atypical reaction
XX pattern of hepatitis markers. PCR amplification and sequencing indicated
XX the 5 specified mutations in the S gene, relative to adw wild-type. These
XX mutations all caused amino acid variations: 4 in the a-determinant region
XX and the other close to it at residue 181. The products of the invention
XX have virucide, hepatotropic and antiinflammatory activity. The products
XX of the invention are used to detect specific antibodies, as probes or
XX primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
XX and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
XX including use in vaccination, diagnosis and screening of donated blood.
XX The antibodies do not bind, or bind much less strongly, to the HBsAg from
XX virus of genotype A, subtype adw. This sequence represents a fragment of
XX Hepatitis B virus surface protein (HBsAg) corresponding to residues 111-
XX 170.
XX SQ Sequence 60 AA;

Alignment Scores:
Pred. No.: 1.248-10 Length: 60
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68793 (1-60)
QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
Db 15 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 34
QY 61 GGAAACTGC 69
Db 35 GlyAsnCys 37

RESULT 14
ADM68806
ID ADM68806 standard; protein; 70 AA.
XX
XX AC ADM68806;
XX
XX DT 24-MAR-2005 (first entry)
XX
XX DE Hepatitis B virus genotype A HBsAg fragment.
XX
XX KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
XX antiinflammatory; DNA detection; vaccine; diagnostic;
XX KW hepatitis B virus infection; selectable marker; a-determinant.
XX
XX OS Hepatitis B virus.
XX
XX PN WO2004113369-A1.
XX
XX PD 29-DEC-2004.
XX
XX PF 17-JUN-2004; 2004WO-BF006515.
XX
XX PR 20-JUN-2003; 2003DB-01028080.
XX
XX PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
XX PI Krupka U;
XX
XX DR WPI; 2005-066451/07.
XX
XX PT New oligo- or poly-peptide variants of hepatitis B surface antigen,
XX useful for diagnosis of, and vaccination against, hepatitis infection,
XX derived from strain HBD05, also related nucleic acid and antibodies.

XX PS Disclosure; Fig 1; 61pp; German.
XX
XX This sequence represents novel oligo- and poly-peptides that are variants
XX of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated
XX from a patient with inflammation of the liver but an atypical reaction
XX pattern of hepatitis markers. PCR amplification and sequencing indicated
XX the 5 specified mutations in the S gene, relative to adw wild-type. These
XX mutations all caused amino acid variations: 4 in the a-determinant region
XX and the other close to it at residue 181. The products of the invention
XX have virucide, hepatotropic and antiinflammatory activity. The products
XX of the invention are used to detect specific antibodies, as probes or
XX primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens
XX and antibodies (Ab) raised against HBsAg are used for detecting HBsAg
XX including use in vaccination, diagnosis and screening of donated blood.
XX The antibodies do not bind, or bind much less strongly, to the HBsAg from
XX virus of genotype A, subtype adw. This sequence represents a fragment of
XX Hepatitis B virus genotype A surface protein (HBsAg) a-determinant
XX corresponding to residues 101-170.
XX SQ Sequence 70 AA;

Alignment Scores:
Pred. No.: 1.268-10 Length: 70
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 9 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68806 (1-70)
QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTTCCTCTTGTGCTGTACAAACCTTCGGAC 60
Db 25 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 44
QY 61 GGAAACTGC 69
Db 45 GlyAsnCys 47

RESULT 15
ADM68821
ID ADM68821 standard; protein; 70 AA.
XX
XX AC ADM68821;
XX
XX DT 24-MAR-2005 (first entry)
XX
XX DE Hepatitis B virus wild type HDB 05 HBsAg S protein fragment.
XX
XX KW antigen; HBsAg; hepatitis marker; virucide; hepatotropic;
XX antiinflammatory; DNA detection; vaccine; diagnostic;
XX KW hepatitis B virus infection; selectable marker; S protein.
XX
XX OS Hepatitis B virus.
XX
XX PN WO2004113369-A1.
XX
XX PD 29-DEC-2004.
XX
XX PF 17-JUN-2004; 2004WO-BP006515.
XX
XX PR 20-JUN-2003; 2003DE-01028080.
XX
XX PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
XX PI Krupka U;
XX
XX DR WPI; 2005-066451/07.
XX
XX PT New oligo- or poly-peptide variants of hepatitis B surface antigen,
XX useful for diagnosis of, and vaccination against, hepatitis infection,
XX derived from strain HBD05, also related nucleic acid and antibodies.
```


XX
PS
XX

Disclosure; Fig 6; 61pp; German.

This sequence represents novel oligo- and poly-peptides that are variants of the hepatitis B surface antigen (HBsAg). HBV strain HBD05 was isolated from a patient with inflammation of the liver but an atypical reaction pattern of hepatitis markers. PCR amplification and sequencing indicated the 5 specified mutations in the S gene, relative to adw wild-type. These mutations all caused amino acid variations: 4 in the a-determinant region and the other close to it at residue 181. The products of the invention of the invention are used to detect specific antibodies, as probes or primers for detecting hepatitis B virus (HBV) nucleic acid, as immunogens and antibodies (Ab) raised against HBsAg are used for detecting HBsAg including use in vaccination, diagnosis and screening of donated blood. The antibodies do not bind, or bind much less strongly, to the HBsAg from virus of genotype A, subtype adw. This sequence represents a fragment of the Hepatitis B virus wild-type HDB 05 surface protein (HBsAg) S protein corresponding to residues 101-170.

XX .SQ Sequence 70 AA;

Alignment Scores:

Pred. No.:	1.26e-10	Length:	70
Score:	130.00	Matches:	21
Percent Similarity:	95.65%	Conservative:	1
Best Local Similarity:	91.30%	Mismatches:	1
Query Match:	93.53%	Indels:	0
DB:	9	Gaps:	0

US-10-761-006A-1_COPY_527_595 (1-69) x ADM68821 (1-70)

QY	1	ACGACTCTGCTCAAGAACTCTACCTTTCCCTTTGTTGCTGTACAAACCTTCGGAC	60
Db	25	ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp	44
QY	61	GGRAACTGC	69
Db	45	GlyAsnCys	47

Search completed: December 27, 2005, 20:51:59
Job time : 20.262 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 27, 2005, 20:44:58 ; Search time 4.6 Seconds
(without alignments)
2480.270 Million cell updates/sec

Title: US-10-761-006A-1_COPY_527_595

Perfect score: 139
Sequence: 1 ACAGCTCTGCTCCTACAGAAA.....AACCTTCGAGCGAACTGC 69

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-O=/cgn2_1/USPO.spool_p/US10761006/runat_27122005_192901_15775/app.query.fasta_1.1102
-DB=issued_Patents_AA -QPMT=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10761006 @Cgn 1 1 01/runat_27122005_192901_15775 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCRTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	139	100.0	400	2	US-09-719-528A-3
2	139	100.0	400	2	US-10-209-264-3
3	133	95.7	154	2	US-09-193-104-25
4	130	93.5	24	1	US-08-776-585-6
5	130	93.5	24	2	US-09-282-826-3
6	130	93.5	154	2	US-09-193-104-10
7	130	93.5	154	2	US-09-193-104-20
8	130	93.5	154	2	US-09-193-104-21
9	130	93.5	225	6	5436139-4
10	130	93.5	226	6	5196194-21
11	130	93.5	226	6	5436139-5
12	130	93.5	281	2	US-09-721-480-3

13	130	93.5	390	2	US-09-721-480-5	Sequence 5, Appli
14	130	93.5	395	6	5196194-18	Patent No. 5196194
15	130	93.5	423	1	US-08-760-797A-1	Sequence 1, Appli
16	130	93.5	424	1	US-08-760-797A-3	Sequence 3, Appli
17	130	93.5	424	2	US-08-932-929B-1	Sequence 1, Appli
18	130	93.5	424	2	US-08-932-929B-3	Sequence 3, Appli
19	130	93.5	531	2	US-09-721-480-7	Sequence 7, Appli
20	127	91.4	40	6	5196194-8	Patent No. 5196194
21	126	90.6	92	2	US-09-193-104-5	Sequence 5, Appli
22	125	89.9	154	2	US-09-193-104-7	Sequence 7, Appli
23	124	89.2	92	2	US-09-193-104-6	Sequence 6, Appli
24	124	89.2	154	2	US-09-193-104-9	Sequence 9, Appli
25	124	89.2	154	2	US-09-193-104-13	Sequence 13, Appli
26	124	89.2	154	2	US-09-193-104-14	Sequence 14, Appli
27	124	89.2	154	2	US-09-193-104-15	Sequence 15, Appli
28	124	89.2	154	2	US-09-193-104-16	Sequence 16, Appli
29	124	89.2	154	2	US-09-193-104-17	Sequence 17, Appli
30	124	89.2	157	2	US-09-306-420C-20	Sequence 20, Appli
31	123	88.5	226	6	5198348-1	Patent No. 5198348
32	122	87.8	154	2	US-09-193-104-18	Sequence 18, Appli
33	121	87.1	92	2	US-09-193-104-4	Sequence 4, Appli
34	121	87.1	154	2	US-09-193-104-12	Sequence 12, Appli
35	121	87.1	154	2	US-09-193-104-19	Sequence 19, Appli
36	121	87.1	154	2	US-09-193-104-22	Sequence 22, Appli
37	121	87.1	154	2	US-09-193-104-23	Sequence 23, Appli
38	121	87.1	154	2	US-09-193-104-24	Sequence 24, Appli
39	121	87.1	226	1	US-08-378-011A-3	Sequence 3, Appli
40	121	87.1	226	2	US-09-471-573A-2	Sequence 2, Appli
41	121	87.1	226	2	US-09-471-573A-40	Sequence 40, Appli
42	121	87.1	236	1	US-08-378-011A-1	Sequence 1, Appli
43	121	87.1	277	6	5164485-2	Patent No. 5164485
44	117	84.2	154	2	US-09-193-104-8	Sequence 8, Appli
45	116	83.5	228	1	US-08-447-591-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-719-528A-3
; Sequence 3, Application US/09719528A
; Patent No. 6558675
; GENERAL INFORMATION:

APPLICANT: Oon, Chong Jin
Lim, Gek Keow
Zhao, Yi

Chen, Wei Ning

TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
USES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ladass & Parry

STREET: 26 West 61 Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10023

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/719,528A

FILING DATE: 30-Apr-2001

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/SG98/00046

FILING DATE: 19-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Mass, Clifford J.

REGISTRATION NUMBER: 30,086

REFERENCE/DOCKET NUMBER: U-013109-7

TELECOMMUNICATION INFORMATION:

```
/
/ TELEPHONE: (212) 708-1800
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 400 amino acids
/   TYPE: amino acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-719-528A-3

Alignment Scores:
Pred. No.:      1.91e-13      Length:      400
Score:          139.00      Matches:      23
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              2          Gaps:      0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-719-528A-3 (1-400)

QY      1  ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db      299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318
QY      61  GGAAACTGC 69
Db      319 GlyAsnCys 321

RESULT 2
US-10-209-264-3
/ Sequence 3, Application US/10209264
/ Patent No. 6787142
/ GENERAL INFORMATION:
/ APPLICANT: Oon, Chong Jin
/ Lim, Gek Keow
/ Zhao, Yi
/ Chen, Wei Ning
/ TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND
/   USES THEREOF
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ladas & Parry
/ STREET: 26 West 61 Street
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10023
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA: US/10/209,264
/ APPLICATION NUMBER: US/10/209,264
/ FILING DATE: 31-Jul-2002
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/SG98/00046
/ FILING DATE: 19-JAN-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Maas, Clifford J.
/ REGISTRATION NUMBER: 30,086
/ REFERENCE/DOCKET NUMBER: U-013109-7
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 708-1800
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 400 amino acids
/   TYPE: amino acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-209-264-3
```

```
Alignment Scores:
Pred. No.:      1.91e-13      Length:      400
Score:          139.00      Matches:      23
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              2          Gaps:      0

US-10-761-006A-1_COPY_527_595 (1-69) x US-10-209-264-3 (1-400)

QY      1  ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db      299 ThrThrProAlaGlnGlyAsnSerThrPheProSerCysCysThrLysProSerAsp 318
QY      61  GGAAACTGC 69
Db      319 GlyAsnCys 321

RESULT 3
US-09-193-104-25
/ Sequence 25, Application US/09193104A
/ Patent No. 6172193
/ GENERAL INFORMATION:
/ APPLICANT: Primi, Daniele
/ APPLICANT: Fiordalisi, Gianfranco
/ APPLICANT: Pallia, Mario
/ TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
/ TITLE OF INVENTION: Virus
/ FILE REFERENCE: SBD1004US
/ CURRENT APPLICATION NUMBER: US/09/193,104A
/ CURRENT FILING DATE: 1998-11-16
/ EARLIER APPLICATION NUMBER: EP 97830635.5
/ EARLIER FILING DATE: 1997-12-01
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 25
/ LENGTH: 154
/ TYPE: PRT
/ ORGANISM: Hepatitis B virus
US-09-193-104-25

Alignment Scores:
Pred. No.:      1.53e-12      Length:      154
Score:          133.00      Matches:      22
Percent Similarity: 95.65%      Conservative: 0
Best Local Similarity: 95.65%      Mismatches: 1
Query Match:     95.68%      Indels:      0
DB:              2          Gaps:      0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-25 (1-154)

QY      1  ACGACTCCTGCTCAAGAACTCTACGTTTCCCTCTTGTTGCTGTACAAAACCTTCGGAC 60
Db      92  ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProSerAsp 111
QY      61  GGAAACTGC 69
Db      112 GlyAsnCys 114

RESULT 4
US-08-776-585-6
/ Sequence 6, Application US/08776585
/ Patent No. 5932426
/ GENERAL INFORMATION:
/ APPLICANT: BARALLE, Francesco E.
/ APPLICANT: SCODELLER, Eduardo
/ APPLICANT: TISMINEZKY, Sergio
/ TITLE OF INVENTION: MOLECULAR PRESENTING SYSTEM
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 3000 K Street, N.W., Suite 500
```

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/776,585
 FILING DATE: 31-MAR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/EP95/03114
 FILING DATE: 04-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: ISACSON, John P.
 REGISTRATION NUMBER: 33,715
 REFERENCE/DOCKET NUMBER: 69582/106
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-776-585-6

Alignment Scores:
 Pred. No.: 3,79e-12 Length: 24
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservatives: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-08-776-585-6 (1-24)
 QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTTGCTGTACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAACCTGC 69
 Db 22 GlyAsnCys 24
 RESULT 5
 US-09-282-826-3
 ; Sequence 3, Application US/09282826A
 ; Patent No. 6403092
 ; GENERAL INFORMATION:
 ; APPLICANT: Pizzo, Salvatore
 ; APPLICANT: Gron, Hanne
 ; TITLE OF INVENTION: IMMUNE RESPONSE MODULATOR ALPHA-2 MACROGLOBULIN COMPLEX
 ; FILE REFERENCE: 2295-1-001CIP
 ; CURRENT APPLICATION NUMBER: US/09/282,826A
 ; CURRENT FILING DATE: 1999-03-31
 ; EARLIER APPLICATION NUMBER: 09/053,301
 ; EARLIER FILING DATE: 1998-04-01
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 24
 ; TYPE: PRT
 ; ORGANISM: HBSAG
 US-09-282-826-3
 Alignment Scores:
 Pred. No.: 3,79e-12 Length: 24

Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservatives: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x US-09-282-826-3 (1-24)
 QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTTGCTGTACAAACCTTCGGAC 60
 Db 2 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 21

QY 61 GGAACCTGC 69
 Db 22 GlyAsnCys 24

RESULT 6
 US-09-193-104-10
 ; Sequence 10, Application US/09193104A
 ; Patent No. 6172193
 ; GENERAL INFORMATION:
 ; APPLICANT: Primi, Daniele
 ; APPLICANT: Fiordalisi, Gianfranco
 ; APPLICANT: Palla, Mario
 ; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
 ; TITLE OF INVENTION: Virus
 ; FILE REFERENCE: SBD1004US
 ; CURRENT APPLICATION NUMBER: US/09/193,104A
 ; CURRENT FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: EP 97830635.5
 ; EARLIER FILING DATE: 1997-12-01
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: Hepatitis B virus
 US-09-193-104-10

Alignment Scores:
 Pred. No.: 4,51e-12 Length: 154
 Score: 130.00 Matches: 21
 Percent Similarity: 95.65% Conservatives: 1
 Best Local Similarity: 91.30% Mismatches: 1
 Query Match: 93.53% Indels: 0
 DB: 1 Gaps: 0
 US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-10 (1-154)

QY 1 ACGACTCTGCTCAAGAACTCTACGTTCCCTCTGTTGCTGTACAAACCTTCGGAC 60
 Db 92 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 111

QY 61 GGAACCTGC 69
 Db 112 GlyAsnCys 114

RESULT 7
 US-09-193-104-20
 ; Sequence 20, Application US/09193104A
 ; Patent No. 6172193
 ; GENERAL INFORMATION:
 ; APPLICANT: Primi, Daniele
 ; APPLICANT: Fiordalisi, Gianfranco
 ; APPLICANT: Palla, Mario
 ; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
 ; TITLE OF INVENTION: Virus
 ; FILE REFERENCE: SBD1004US
 ; CURRENT APPLICATION NUMBER: US/09/193,104A
 ; CURRENT FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: EP 97830635.5
 ; EARLIER FILING DATE: 1997-12-01
 ; NUMBER OF SEQ ID NOS: 44

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-20

Alignment Scores:
Pred. No.: 4,51e-12 Length: 154
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 91.30% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-20 (1-154)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 92 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 111
QY 61 GGAAACTGC 69
Db 112 GlyAsnCys 114

RESULT 8
US-09-193-104-21
; Sequence 21, Application US/09193104A
; Patent No. 6172193
; GENERAL INFORMATION:
; APPLICANT: Primi, Daniele
; APPLICANT: Fiordalisi, Gianfranco
; APPLICANT: Palla, Mario
; TITLE OF INVENTION: Escape Mutant of the Surface Antigen of Hepatitis B
; FILE REFERENCE: SBD1004US
; CURRENT APPLICATION NUMBER: US/09/193,104A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: EP 97830635.5
; EARLIER FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Hepatitis B virus
US-09-193-104-21

Alignment Scores:
Pred. No.: 4,51e-12 Length: 154
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 91.30% Indels: 0
DB: 2 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-09-193-104-21 (1-154)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 92 ThrThrProAlaGlnGlyAsnSerLeuPheProSerCysCysThrLysProThrAsp 111
QY 61 GGAAACTGC 69
Db 112 GlyAsnCys 114

RESULT 9
5436139-4
; Patent No. 5436139
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: NON-PASSAGEABLE VIRUS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/058/89,993
; FILING DATE: 08-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 571,331
; FILING DATE: 22-AUG-1990
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 4
; LENGTH: 225
5436139-4

Alignment Scores:
Pred. No.: 4,67e-12 Length: 225
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 91.30% Indels: 0
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x 5436139-4 (1-225)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 124 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 143
QY 61 GGAAACTGC 69
Db 144 GlyAsnCys 146

RESULT 10
5196194-21
; Patent No. 5196194
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,621
; FILING DATE: 7-DEC-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 21
; LENGTH: 226
5196194-21

Alignment Scores:
Pred. No.: 4,67e-12 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 91.30% Indels: 0
DB: 6 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x 5196194-21 (1-226)

QY 1 ACGACTCCTGCTCAAGGAACTCTACGTTCCCTCTTGTGCTGTACAAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrLysProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 11
5436139-5
```

Patent No. 5436139
; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: NON-PASSAGEABLE VIRUS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/058/89,993
; FILING DATE: 08-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 571,331
; FILING DATE: 22-AUG-1990
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 5:
; LENGTH: 226
5436139-5

Alignment Scores:
Pred. No.: 4.67e-12 Length: 226
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 6 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x 5436139-5 (1-226)

QY 1 ACAGCTCTGCTCAAGAACTCTAGTTTCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrlYsProThrAsp 144
QY 61 GGAAACTGC 69
Db 145 GlyAsnCys 147

RESULT 12

US-09-721-480-3
; Sequence 8, Application US/09721480
; Patent No. 6740323

; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-721-480-3

Alignment Scores:
Pred. No.: 4.77e-12 Length: 281
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x US-09-721-480-3 (1-281)

QY 1 ACAGCTCTGCTCAAGAACTCTAGTTTCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 125 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrlYsProThrAsp 144

Db 180 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrlYsProThrAsp 199
QY 61 GGAAACTGC 69
Db 200 GlyAsnCys 202
RESULT 13
US-09-721-480-5
; Sequence 9, Application US/09721480
; Patent No. 6740323
; GENERAL INFORMATION:
; APPLICANT: Selby, Mark
; APPLICANT: Glazer, Edward
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE
; FILE REFERENCE: PP01635.002
; CURRENT APPLICATION NUMBER: US/09/721,480
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCMVII opti
US-09-721-480-5

Alignment Scores:
Pred. No.: 4.92e-12 Length: 390
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 2 Gaps: 0

US-10-761-006a-1_COPY_527_595 (1-69) x US-09-721-480-5 (1-390)

QY 1 ACAGCTCTGCTCAAGAACTCTAGTTTCCTCTGTGCTGTACAAACCTTCGGAC 60
Db 289 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysThrlYsProThrAsp 308
QY 61 GGAAACTGC 69
Db 309 GlyAsnCys 311

RESULT 14

5196194-18
; Patent No. 5196194

; APPLICANT: RUTTER, WILLIAM J.; GOODMAN, HOWARD M.
; TITLE OF INVENTION: VACCINES CONTAINING HEPATITIS B S-PROTEIN
; NUMBER OF SEQUENCES: 21
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,621
; FILING DATE: 7-DEC-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 513,055
; FILING DATE: 12-JUL-1983
; APPLICATION NUMBER: 107,267
; FILING DATE: 21-DEC-1979
; APPLICATION NUMBER: 41,909
; FILING DATE: 24-MAY-1979
; SEQ ID NO: 18
; LENGTH: 395
5196194-18

Alignment Scores:
Pred. No.: 4.92e-12 Length: 395
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0

DB: 6 Gaps: 0
US-10-761-006A-1_COPY_527_595 (1-69) x 5196194-18 (1-395)
QY 1 ACAGACTCCGCTCAAGAACTCTAGCTTCCCTCTTGTTGTGTACAAAACCTTCGGAC 60
Db 294 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 313
QY 61 GGAAACTGC 69
Db 314 GlyAsnCys 316

Db 342 GlyAsnCys 344

Search completed: December 27, 2005, 21:03:41
Job time : 4.6 secs

1026

RESULT 15
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5935902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: Hybrid Protein Between CS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-760-797A-1

Alignment Scores:
Pred. No.: 4.95e-12 Length: 423
Score: 130.00 Matches: 21
Percent Similarity: 95.65% Conservative: 1
Best Local Similarity: 91.30% Mismatches: 1
Query Match: 93.53% Indels: 0
DB: 1 Gaps: 0

US-10-761-006A-1_COPY_527_595 (1-69) x US-08-760-797A-1 (1-423)

QY 1 ACAGACTCCGCTCAAGAACTCTAGCTTCCCTCTTGTTGTGTACAAAACCTTCGGAC 60
Db 322 ThrThrProAlaGlnGlyAsnSerMetPheProSerCysCysCysThrLysProThrAsp 341
QY 61 GGAAACTGC 69
|||||